Sujay V. Kumar

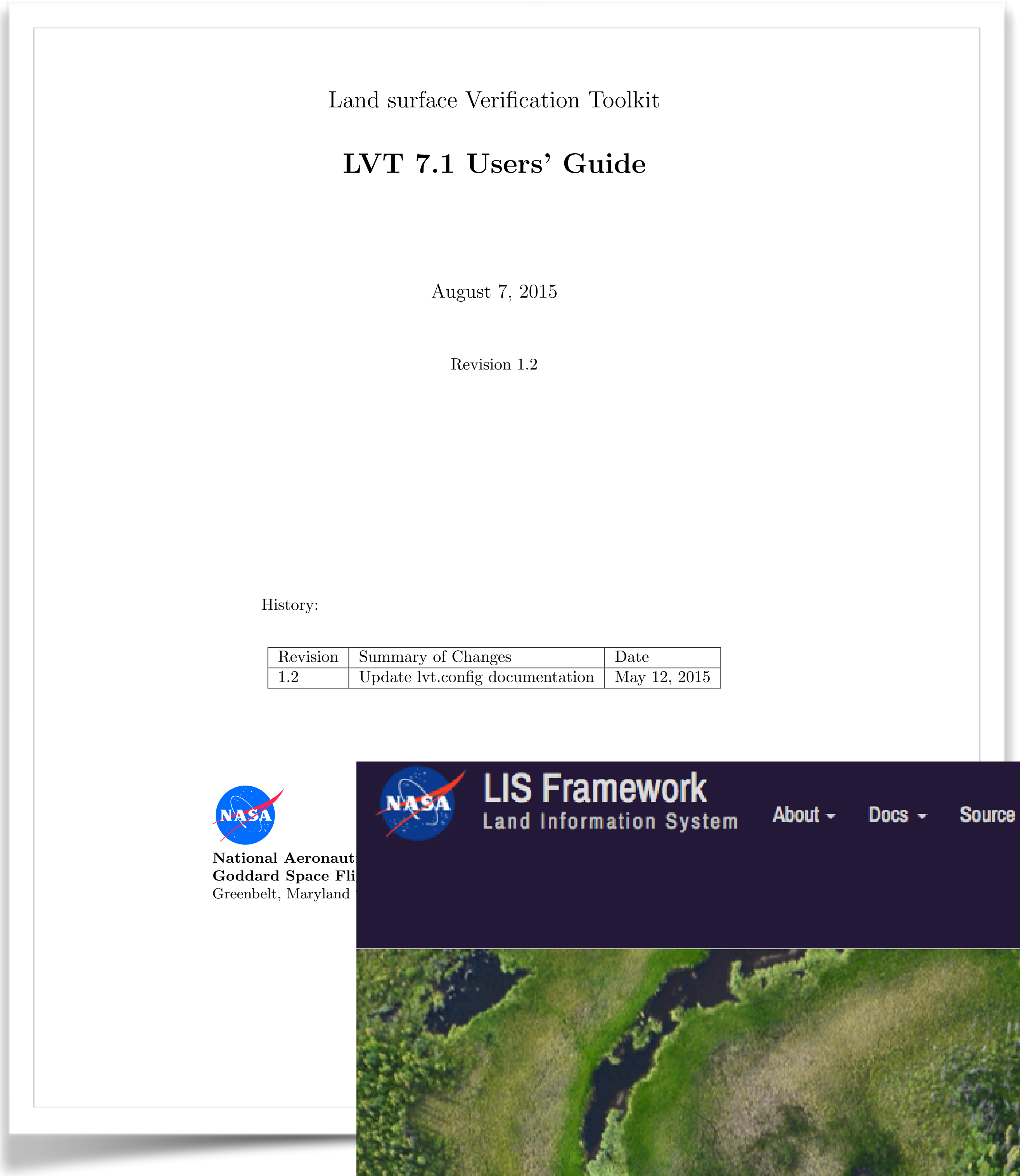
Science Applications International Corporation/NASA Hydrological Sciences Laboratory

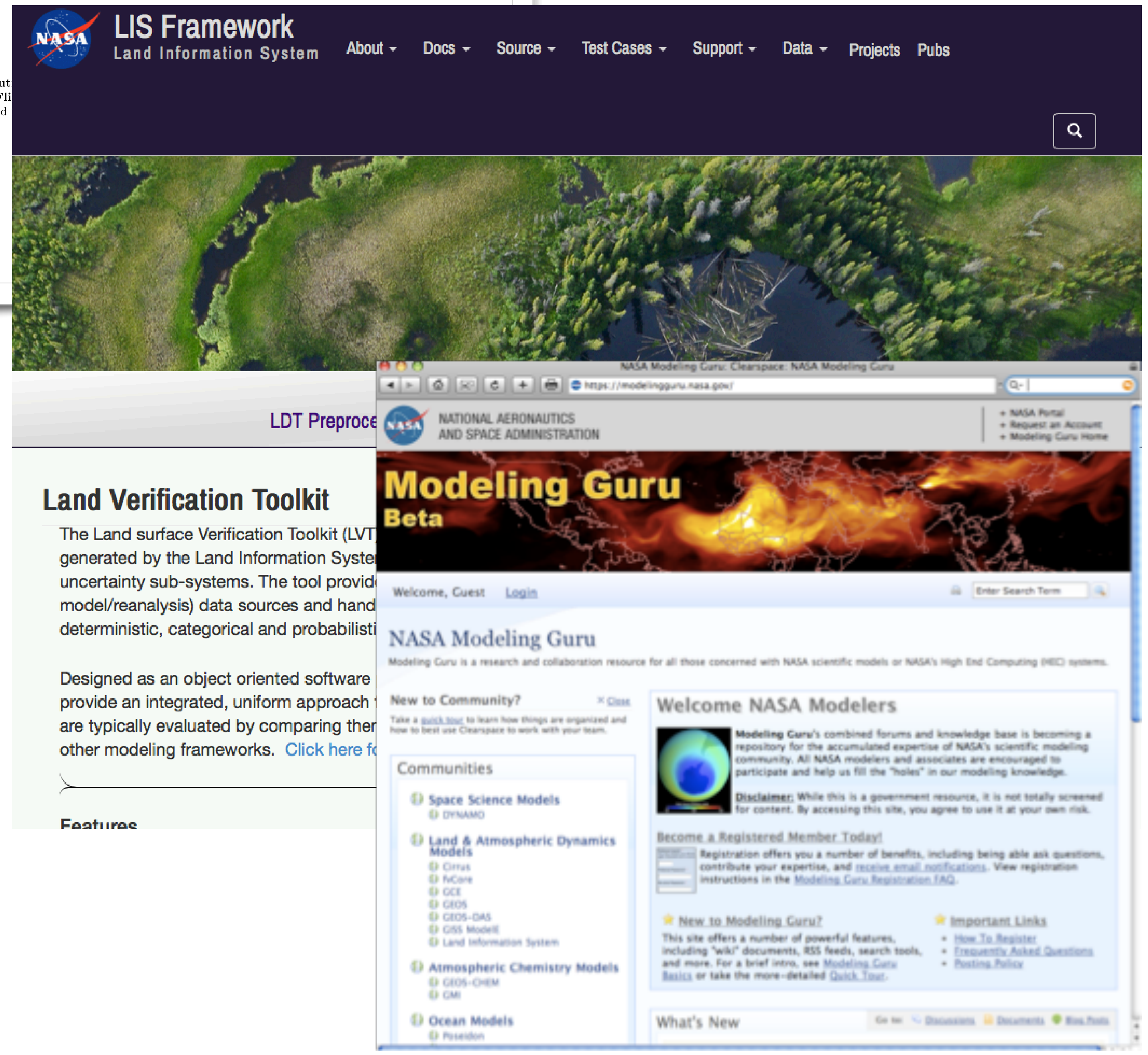
Software requirements

- ☑ Fortran 90/95 compiler (intel, gfortran preferred)
- ☑ C compiler
- ☑ Earth System Modeling Framework (ESMF; 5.x or greater) (https://www.earthsystemcog.org/projects/esmf/)
- ☑ NetCDF (3 or 4; http://www.unidata.ucar.edu/software/netcdf/)
- ☑ Grib-API (https://software.ecmwf.int/wiki/display/GRIB/Home)
- ☐ HDF5 (https://www.hdfgroup.org/HDF5/)
- ☐ HDF4/HDFEOS

Software maintenance

- Subversion repository (https://progress.nccs.nasa.gov)
  - Accessible to NCCS users
- User's guide
  - Step-by-step instructions on how to build the LVT code
- Reference manual/in-line documentation
- http://lis.gsfc.nasa.gov
- http://modelingguru.nasa.gov





Building LVT

- Build the required software libraries
- Setup the LVT environment variables

Variable	Description
LVT_SRC	Location of the LVT source tree (*$WORKING/src/*)
LVT_ARCH	LVT architecture (See below)
LVT_FC	Fortran compiler to be used (*mpif90*, if mpi is installed)
LVT_CC	C compiler to be used (*mpicc*, if mpi is installed)
LVT_GRIBAPI	path to grib api library
LVT_NETCDF	path to NETCDF library
LVT_HDF4	path to HDF4 library
LVT_HDF5	path to HDF5 library
LVT_HDFEOS	path to HDFEOS library
LVT_MODESMF	path to ESMF header files
LVT_LIBESMF	path to ESMF library files

- Run the configure script, followed by the compile script.

```
~/LVTv1.0/src % ./configure
-----------------------------------------------------------------------
Setting up configuration for LVT version 1.0...
Optimization level (-2=strict checks, -1=debug, 0,1,2,3, default=2):
Assume little/big_endian data format (1-little, 2-big, default=2):
Use NETCDF? (1-yes, 0-no, default=1):
NETCDF version (3 or 4, default=4):
NETCDF use shuffle filter? (1-yes, 0-no, default = 1):
NETCDF use deflate filter? (1-yes, 0-no, default = 1):
NETCDF use deflate level? (1 to 9-yes, 0-no, default = 9):
Use HDF4? (1-yes, 0-no, default=1):
Use HDF5? (1-yes, 0-no, default=1):
Use HDFEOS? (1-yes, 0-no, default=1):
----------------------------------------------------
 configure.lvt file generated successfully
----------------------------------------------------
Settings are written to configure.lvt in the make directory
If you wish to change settings, please edit that file.
To compile, run the compile script.
-----------------------------------------------------------------------
~/LVTv1.0/src % ./compile
-----------------------------------------------------------------------
Compiling LVT version 1.0...
-----------------------------------------------------------------------
Building dependency generator...
/usr/local/intel/Composer/composer_xe_2013_sp1.3.174/bin/intel64/icc -c -O main.c
/usr/local/intel/Composer/composer_xe_2013_sp1.3.174/bin/intel64/icc -o makdep main.o
-----------------------------------------------------------------------
Compiling LVT source code...
Building dependency file vinterp.d
Building dependency file upscaleByAveraging_input.d
Building dependency file upscaleByAveraging.d
Building dependency file template_obsMod.d
Building dependency file stninterp_module.d
Building dependency file readtemplateObs.d
Building dependency file readinput_latlon.d
Building dependency file readinput_lambert.d
Building dependency file readinput_UTM.d
Building dependency file read_LSWG_Tb_Obs.d
Building dependency file readWGPBMRObs.d
```

```
sp3/lib/libO/Linux.intel.64.intelmpi.default -lesmf -lstdc++ -limf -lm -lrt -L/discover/nobackup/projects/lis/
libs/grib_api/1.12.3_intel-14.0.3.174_sp3/lib/ -lgrib_api_f90 -lgrib_api -L/discover/nobackup/projects/lis/lib
s/jasper/1.900.1_intel-14.0.3.174_sp3/lib/ -ljasper -L/discover/nobackup/projects/lis/libs/netcdf/4.3.3.1_inte
l-14.0.3.174_sp3/lib/ -lnetcdff -lnetcdf -L/discover/nobackup/projects/lis/libs/hdfeos2/2.19v1.00_intel-14.0.3
.174_sp3/lib/ -lhdfeos -lGctp -L/discover/nobackup/projects/lis/libs/hdf4/4.2.11_intel-14.0.3.174_sp3/lib/ -lm
fhdf -ldf -ljpeg -lz  -L/discover/nobackup/projects/lis/libs/hdf5/1.8.14_intel-14.0.3.174_sp3/lib/ -lhdf5_fort
ran -lhdf5_hl -lhdf5
Compile finished
-----------------------------------------------------------------------
~/LVTv1.0/src % ls -l LVT
-rwx------ 1 svkumar s1189 22873207 2015-08-20 14:00 LVT*
~/LVTv1.0/src %
```

Example 1

LIS Noah LSM output vs USDA ARS in-situ surface soil moisture measurements

LVT configuration

- For the text entries, case/exactness of the string is important!
- For entries with spaces, use double quotes. Otherwise quotes are not necessary
- Comments can be inserted with a # prefix

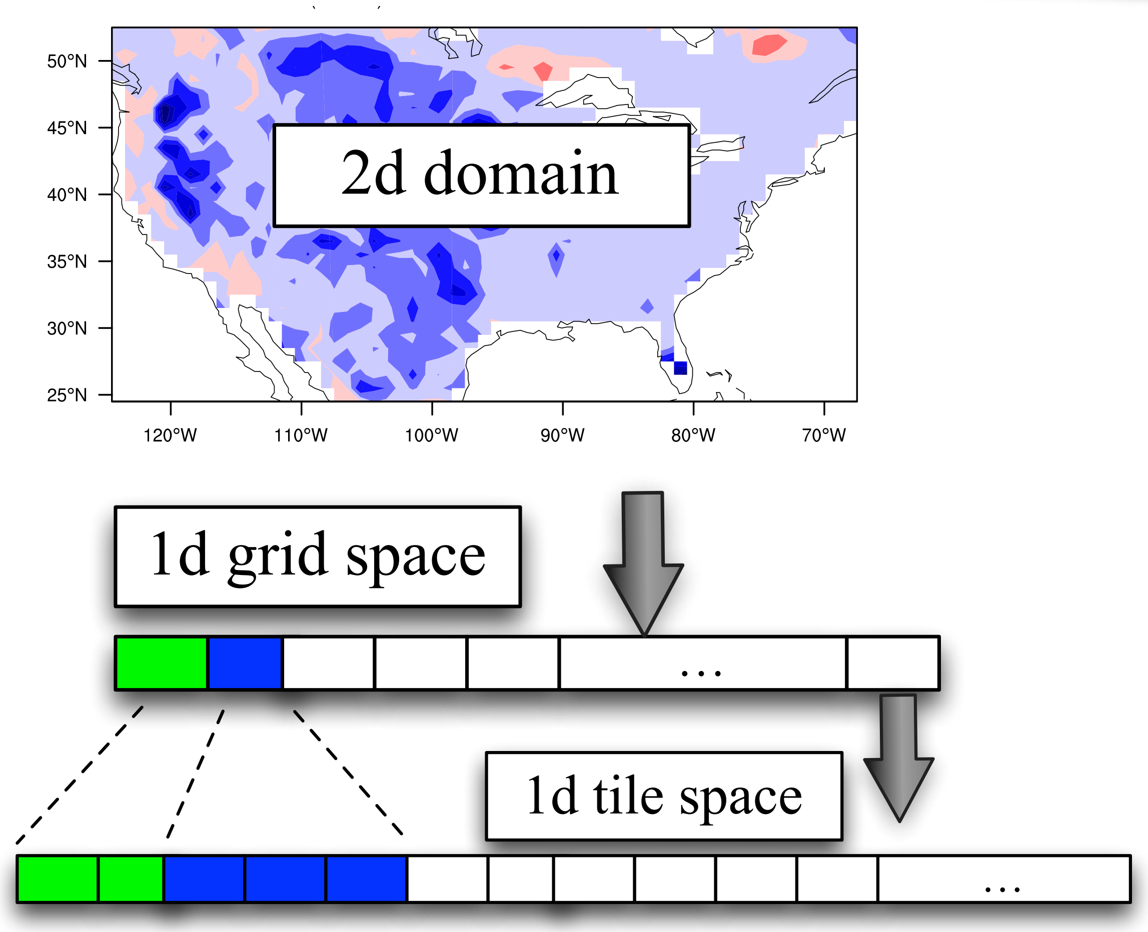
```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from a
# LIS output (from Noah.3.3 LSM) against the in-situ USDA ARS soil moisture
# measurements
#
# The model output from Noah.3.3 output is produced over CONUS at 0.125 deg
# spatial resolution (at daily intervals). The LVT analysis is conducted
# over a CONUS domain at 0.5 deg spatial resolution.
#
# The following variables are compared: surface soil moisture
#
# The following metrics are used: Mean and Anomaly correlation
#
#------------------------------------------------------------------------------

LVT running mode:                      "Data intercomparison"
Map projection of the LVT analysis:    "latlon"
LVT output format:                     "netcdf"
LVT output methodology:                "2d gridspace"
Analysis data sources:                 "LIS output"    "USDA ARS soil moisture"
```

LVT running mode: specifies the running mode to be used Acceptable values are:

Value	Description
"Data intercomparison"	standard analysis mode where a particular data is compared against another
"Benchmarking"	A benchmarking output is generated based on the input training datasets

LVT output methodology: specifies the output methodology used in LVT. The LVT output is written as a 1-D array containing only land points or as a 2-D array containing both land and water points. 1-d tile space includes the subgrid tiles and ensembles. 1-d grid space includes a vectorized, land-only grid-averaged set of values. Acceptable values are:

Value	Description
"1d tilespace"	LVT output in a 1-D tile domain
"2d gridspace"	LVT output in a 2-D grid domain
"1d gridspace"	LVT output in a 1-D grid domain



Configuration: Data sources

```
#--------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from a
# LIS output (from Noah.3.3 LSM) against the in-situ USDA ARS soil moisture
# measurements
#
# The model output from Noah.3.3 output is produced over CONUS at 0.125 deg
# spatial resolution (at daily intervals). The LVT analysis is conducted
# over a CONUS domain at 0.5 deg spatial resolution.
#
# The following variables are compared: surface soil moisture
#
# The following metrics are used: Mean and Anomaly correlation
#
#--------------------------------------------------------------------------

LVT running mode:                   "Data intercomparison"
Map projection of the LVT analysis: "latlon"
LVT output format:                  "netcdf"
LVT output methodology:             "2d gridspace"
Analysis data sources:              "LIS output" "USDA ARS soil moisture"
```

`Analysis data sources:` specifies the two sources of data to be used in an LVT comparison. The user must always choose two sources specified in adjacent columns. The second column entry will be taken as the reference data and the first column will be used as the data being evaluated (against the reference data). If the comparison LIS output against a non-LIS data, it is recommended to specify the first source as "LIS output" and then the other data as the second data source.

- Two data sources must always be specified, separated by spaces

  "LIS output" "LIS output"

  "LIS output" "none"

  "LIS output" "NLDAS2"

  "NLDAS2" "CMC"

  "NLDAS2" "GDAS2"

Value	Description
"none"	template
"LIS output"	output from another LIS run
"LIS DAOBS"	processed observations from a LIS DA run
"ISCCP LST"	ISCCP skin temperature observations
"MODIS LST"	MODIS (Terra/Aqua) land surface temperature observations
"SCAN"	SCAN soil moisture station observations
"NASMD"	North American Soil Moisture Database soil moisture station
"ISMN"	ISMN soil moisture station observations
"SURFRAD"	SURFRAD observations
"SNOTEL"	SNOTEL snow water equivalent observations
"LSWG Tb"	Tb brightness temperature observations at the LSWG sites
"FMI SWE"	Finnish Meteorological Institute (FMI) snow course data
"CMC"	Canadian Meteorological Center (CMC) snow depth analysis
"SNODAS"	NOHRSC SNow Data Assimilation (SNODAS) product
"AMSR-E NASA soil moisture"	NASA (NSIDC) retrieval of AMSR-E soil moisture
"AMSR-E LPRM soil moisture"	LPRM (VU) retrieval of AMSR-E soil moisture
"AMMA"	AMMA station observations
"Ameriflux"	Ameriflux station observations
"ARM"	ARM station observations
"SMOSREX"	SMOSREX station observations
"AGRMET"	AGRMET land surface analysis
"Globsnow"	GlobSnow SWE analysis
"SNODEP metobs"	WMO snow depth station observations
"MOD10A1"	MOD10A1 fractional snow cover data from MODIS
"ANSA snowdepth"	ANSA snow depth retrievals
"ANSA SWE"	ANSA SWE retrievals
"CPC precipitation"	CPC unified precipitation product
"USGS streamflow"	USGS streamflow observations
"Naturalized streamflow"	Naturalized streamflow estimates
"FLUXNET"	Gridded FLUXNET data from MPI
"MOD16A2"	MOD16A2 ET products from MODIS
"UW ET"	University of Washington ET products from MODIS
"ALEXI"	ALEXI model ET estimates from USDA
"USDA ARS soil moisture"	soil moisture measurements from USDA ARS watersheds
"GHCN"	Global Historical Climatology Network data
"ALEXI"	Atmosphere Land Exchange Inverse model outputs of ET
"NLDAS2"	North American Land Data Assimilation System Phase-2 data
"GRACE"	processed GRACE data used in a LIS-DA instance
"PBO H2O"	plate boundary observatory data
"USGS ground water well data"	USGS ground water well data
"SMOS L2 soil moisture"	SMOS level 2 soil moisture
"SMOS L1 TB"	SMOS level 1 brightness temperature
"GCOMW AMSR2 L3 soil moisture"	GCOMW AMSR2 level 3 soil moisture
"SMOPS soil moisture"	Soil Moisture Operational Product System data
"ESA CCI soil moisture"	ESA CCI soil moisture
"GIMMS NDVI"	GIMMS NDVI data
"GLDAS2"	NASA Global Land Data Assimilation System version 2 data
"MERRA2"	MERRA version 2 data
"GLERL hydro data"	Great Lakes hydrology data
"GL6 JULES data"	GL6 JULES data

Configuration: Time specification

```
Start mode:                         coldstart
LVT restart output interval:        "1mo"
LVT restart filename:               none
Starting year:                      2006
Starting month:                     5
Starting day:                       1
Starting hour:                      0
Starting minute:                    0
Starting second:                    0
Ending year:                        2006
Ending month:                       8
Ending day:                         31
Ending hour:                        0
Ending minute:                      0
Ending second:                      0
LVT clock timestep:                "1da"
Undefined value:                   -9999
LVT diagnostic file:                lvtlog
```

`LVT restart output interval`: specifies the frequency at which the restart files must be written during a LVT analysis. The time interval is specified with a number followed by a 2 character suffix that indicates the units. For example, a restart interval of 1 hour can be specified as “1hr”, “60mn”, or “3600ss”.

Acceptable values for the 2 character suffix are:

Value	Description
ss	second
mn	minute
hr	hour
da	day
mo	month
yr	year

LVT recomputes the clock timestep based on the data intervals of each data stream. The minimum tilmestep value is chosen.

LVT clock timestep = min (timestep set in the config file, timestep of datastream 1, timestep of datastream 2)

Configuration: Analysis domain

```
Input domain and mask data file:   ./lis_input.nldas.nc
#LVT analysis domain
Run domain lower left lat:              25.75
Run domain lower left lon:            -124.75
Run domain upper right lat:             50.75
Run domain upper right lon:            -67.75
Run domain resolution (dx):             0.5
Run domain resolution (dy):             0.5
```

- Specifies the extents of the LVT analysis domain
- The config entries are dependent on the chosen LVT map projection
- The LVT analysis domain can be a subset of the domain specified in the 'Input domain and mask data file'
- The spatial resolution of the LVT analysis domain can be different from the spatial resolution of the 'Input domain and mask data file'
- LVT will generate the landmask for the analysis domain by interpolating/upscaling the 'LANDMASK' field

The input NetCDF file used to create the gridspace in LVT

Should contain a field called "LANDMASK" with a 0/1 landmask representation

The global attributes/dimensions of this file should contain relevant map projection and domain extent information

```
netcdf lis_input.nldas {
dimensions:
        east_west = 464 ;
        north_south = 224 ;

variables:
        float time(time) ;
        float LANDMASK(north_south, east_west) ;
                LANDMASK:standard_name = "LANDMASK" ;
                LANDMASK:units = "" ;
                LANDMASK:scale_factor = 1.f ;
                LANDMASK:add_offset = 0.f ;
                LANDMASK:missing_value = -9999.f ;
                LANDMASK:vmin = 0.f ;
                LANDMASK:vmax = 0.f ;
                LANDMASK:num_bins = 1 ;

 // global attributes:
                :MAP_PROJECTION = "EQUIDISTANT CYLINDRICAL" ;
                :SOUTH_WEST_CORNER_LAT = 25.0625f ;
                :SOUTH_WEST_CORNER_LON = -124.9375f ;
                :DX = 0.125f ;
                :DY = 0.125f ;
```

Configuration: Datastream attributes table

```
#model soil moisture vs obs soil moisture
LVT datastream attributes table::
SoilMoist     1  m3/m3    -     1   4 SoilMoist 1  m3/m3    -    1  1
::
```

Variable from datastream 1 | Variable from datastream 2

- Variable name
- Number of selected levels (0=>not selected)
- Units
- Direction type
- Time averaging option (0-instantaneous, 1- time averaged)
- Number of total vertical levels

```
LVT datastream attributes table::
Qle          1  W/m2    UP   1   1 SoilMoist   1  m3/m3   -   1  4
Qle          1  W/m2    UP   1   1 TotalPrecip 1  kg/m2   -   1  1
Qle          1  W/m2    UP   1   1 Rnet        1  W/m2    -   1  1
RadT         0  K       -    1   1 RadT        0  m3/m3   -   1  4
::
```

- Specifies the variables being analyzed from the datastreams
- Each line represents variable specification from datastream 1 and datastream 2
- Specification for each variable consists of 6 columns; variable names follow ALMA convention
- Any variable from datastream 1 can be compared to any variable from datastream 2 (as long as the metric of comparison is meaningful!)

- Qle being compared against surface soil moisture, total precipitation, Net radiation
- Qle from datastream 1 is used for multiple comparisons against variables from datastream 2
- Radiative temperature comparison is turned off

Configuration: Vertical averaging, external masking

```
LVT surface soil layer thickness:  0.10
LVT root zone soil layer thickness:  1.0
```

- Specifies the thickness (m) of the surface and root zone soil layers used in the analysis
- LVT will vertically average the individual soil layers (of the datastreams) to the thickness used in analysis
- The averaging will be weighted by the thicknesses

```
Apply external mask:                 0
External mask directory:             none
Temporal (monthly) mask flags:       0 0 0 0 0 1 1 1 0 0 0 0
```

- 0 - no masking
- 1- external temporally varying mask
- 2 - external fixed static mask
- 3 - temporal monthly mask

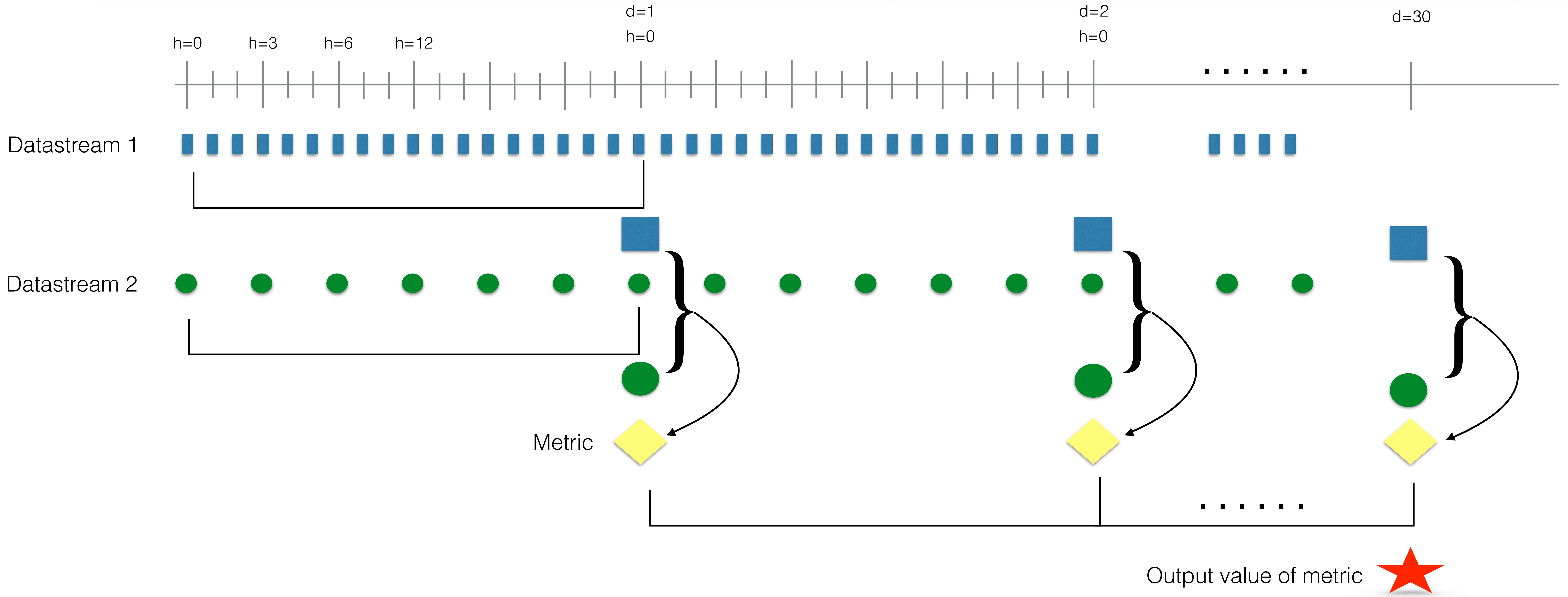
External mask directory: Specifies the name of the data mask file/directory. If the mask varies temporally, then this option specifies the top-level directory containing data mask. Note that the mask files are expected to be in binary, sequential access format.

- Apply monthly mask (0/1) in the analysis (restricts the analysis to JJA in the above case)

Configuration: Output frequency

```
Observation count threshold:                              0
Temporal averaging interval:                              "1da"
Stats output interval:                                    "1mo"
Starting month if a shifted year definition is used in temporal averaging: 1
```

- Observation count threshold - computations are excluded over those grid points where the specified minimum count is not met
- Temporal averaging interval - the individual datastream values are averaged upto this interval and then the metric is computed
- Stats output interval - the metric values are averaged upto the stats output interval



Configuration: Time series output

- Specifies the name of the file that lists the locations and regions in the domain where ASCII time series data (for each metric) are to be derived
- Five different styles of specifying the locations/regions

```
Time series location file:                    ./TS_LOCATIONS.TXT
```

Style 1: Specify lat/lon bounding boxes

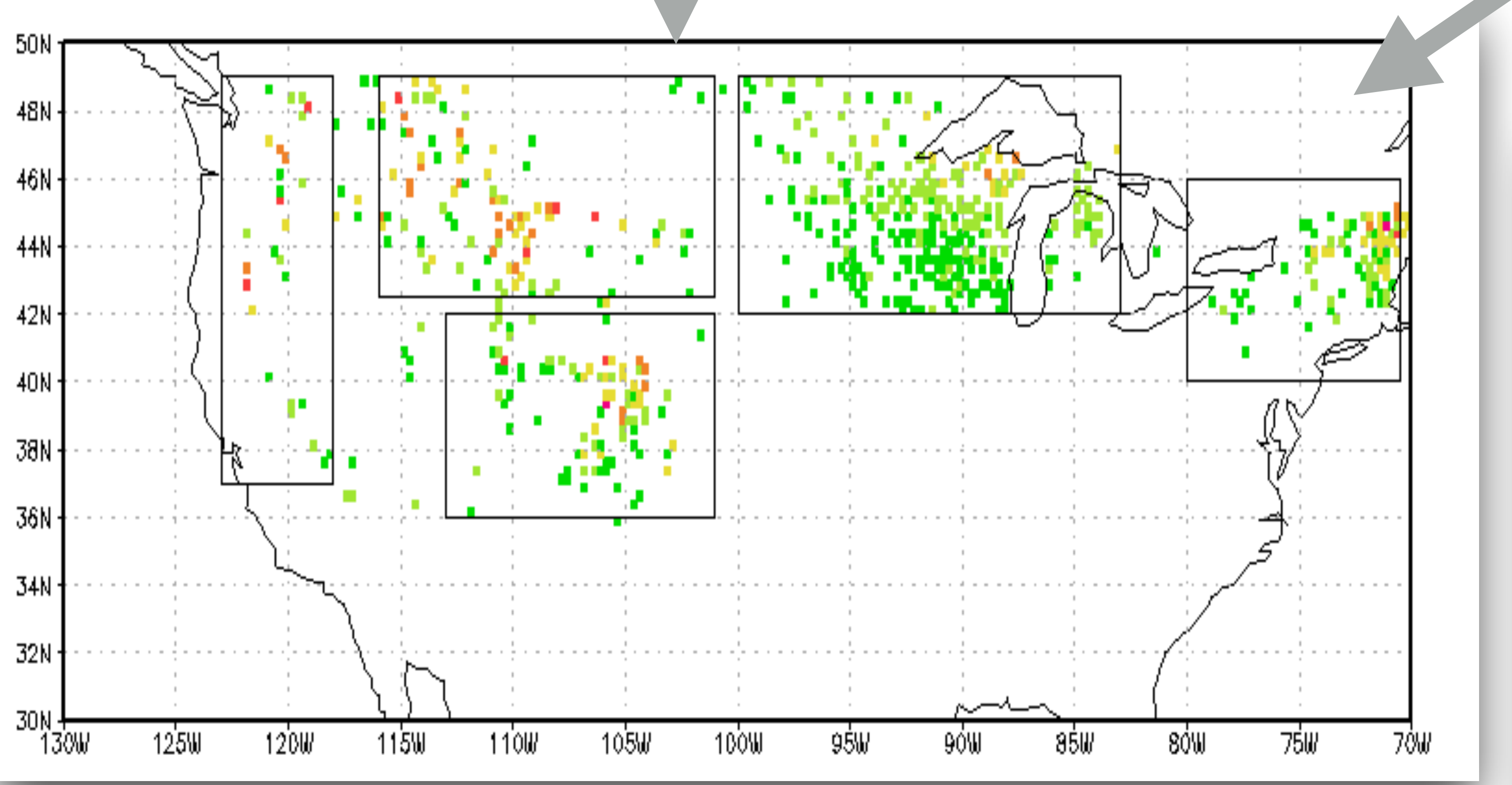
```
#Number of locations
2
#Location style (1-lat/lon, 2-col/row, 3-tile)
1
#Location name, (next line) SW-lat, SW-lon, NE-lat, NE-lon,
min number of grid points
WEST_US
40 -130 50 -110 5
HIGH_PLAINS_US
43 -110 49 -100 2
 .....
 .....
```

Style 2: Specify column/row bounding indices

```
#Number of locations
1
#Location style (1-lat/lon, 2-col/row, 3-tile)
2
#Location name, (next line) SW-col, SW-row, NE-col, NE-row,
min number of grid points
WEST_US
1  1  20 30 5
EAST_US
1  1  10 10 5
 .....
 .....
```

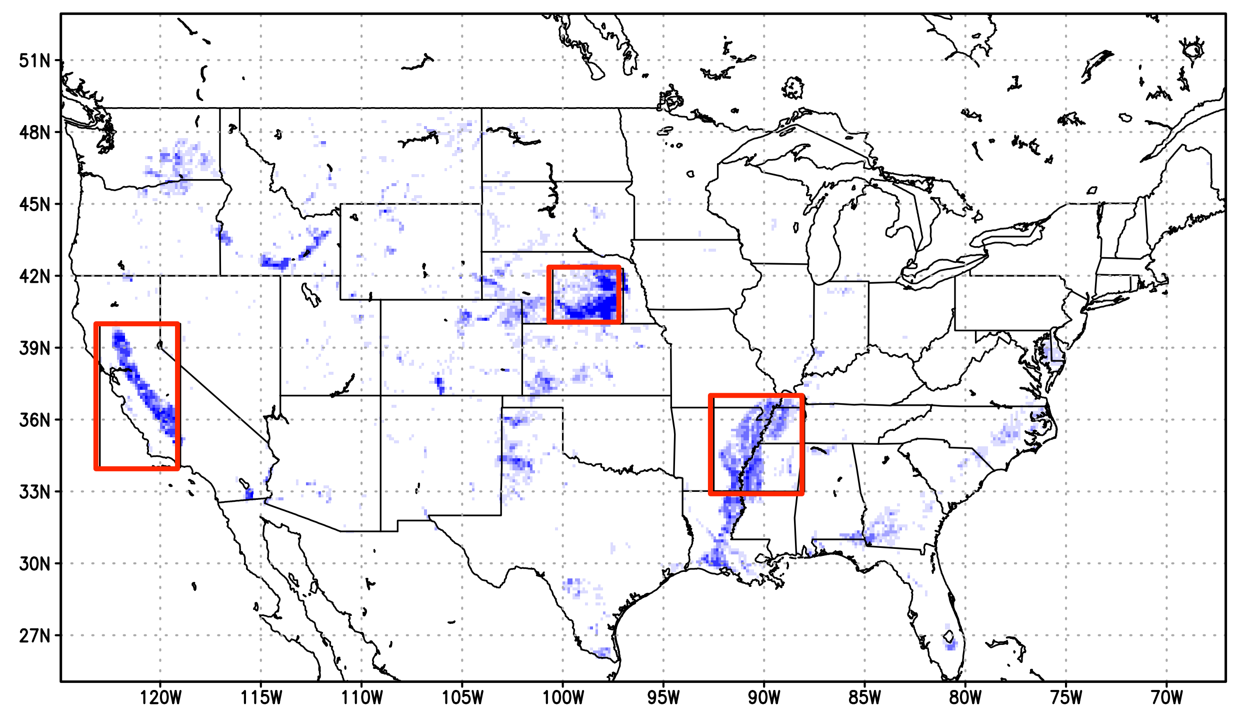
Style 3: Specify 1-d tile bounding indices

```
#Number of locations
1
#Location style (1-lat/lon, 2-col/row, 3-tile)
3
#Location name, (next line) Start index, Ending index,
min number of grid points
WEST_US
1  20 5
EAST_US
1  10 5
.....
.....
```



Style 4: List of lat/lons to specify a region

```
#Number of locations
2
#Location style
4
#number of points followed by lat/lon of each point
REGION1
3
34.4 -103.2
33.4 -100.2
32.1 -99.3
REGION2
2
40.2 -103.3
42.2 -104.2
```



Style 5: Regions defined by a categorical map

```
#Number of stations
3
#style
5
#names
NEWENGLAND
1
MIDATLANTIC
1
SOUTHATLANTIC
1
#categorical map
../huc02_conus_0.125dg.1gd4r
```

Configuration: Metrics attributes

```
Metrics attributes file:                    ./METRICS.TBL
```

```
#name                                   total in-time writeTS extractTS threshold ASC ADC
Mean:                                     1     1      0     1    -9999.0   0   0   #Mean
Min:                                      0     0      0     0    -9999.0   0   0   #Minimum
MinTime:                                  0     0      0     0    -9999.0   0   0   #MinTime
Max:                                      0     0      0     0    -9999.0   0   0   #Maximum
MaxTime:                                  0     0      0     0    -9999.0   0   0   #MaxTime
Sum:                                      0     0      0     0    -9999.0   0   0   #Sum
Anomaly:                                  0     0      0     0    -9999.0   0   0   #Anomaly
Standard deviation:                       0     0      0     0    -9999.0   0   0   #Std
RMSE:                                     0     0      0     0    -9999.0   0   0   #RMSE
Bias:                                     0     0      0     0    -9999.0   0   0   #Bias
ubRMSE:                                   0     0      0     0    -9999.0   0   0   #ubRMSE
Mean absolute error:                      0     0      0     0    -9999.0   0   0   #MAE
Anomaly RMSE:                             0     0      0     0    -9999.0   0   0   #ARMSE
Anomaly correlation:                      0     0      0     0    -9999.0   0   0   #ARMSE
Raw correlation:                          0     0      0     0    -9999.0   0   0   #RCORR
Probability of detection (PODy):          0     0      0     0    0.1       0   0   #PODy
Probability of detection (PODn):          0     0      0     0    0.1       0   0   #PODn
False alarm ratio (FAR):                  0     0      0     0    0.1       0   0   #FAR
Probability of false detection (POFD):    0     0      0     0    0.1       0   0   #POFD
Critical success index (CSI):             0     0      0     0    0.1       0   0   #CSI
Accuracy measure (ACC):                   0     0      0     0    0.1       0   0   #ACC
Frequency bias (FBIAS):                   0     0      0     0    0.1       0   0   #FBIAS
Equitable threat score (ETS):             0     0      0     0    0.1       0   0   #ETS
Area metric:                              0     0      0     0    -9999.0   0   0   #Area
Nash sutcliffe efficiency:                0     0      0     0    -9999.0   0   0   #NSE
Ensemble mean:                            1     1      0     1    -9999.0   0   0   #ensmean

Ensemble standard deviation:              1     1      0     1    -9999.0   0   0   #ensstd
Ensemble likelihood:                      1     1      0     1    -9999.0   0   0   #ensll
Ensemble cross correlation:               1     1      0     1    -9999.0   0   0   #ensxcorr
Ensemble skill:                           0     0      0     0    -9999.0   0   0   #ensskill
Ensemble mean error:                      0     0      0     0    -9999.0   0   0   #ensmerror
Ensemble mean bias:                       0     0      0     0    -9999.0   0   0   #ensmbias
Ensemble spread:                          0     0      0     0    -9999.0   0   0   #ensspread
Metric entropy:                           0     0      0     0    -9999.0   0   0   #mentropy
Information gain:                         0     0      0     0    -9999.0   0   0   #igain
Fluctuation complexity:                   0     0      0     0    -9999.0   0   0   #fcomplexity
Effective complexity:                     0     0      0     0    -9999.0   0   0   #ecomplexity
Wavelet stat:                             0     0      0     0    -9999.0   0   0   #waveletstat
Hausdorff norm:                           0     0      0     0    -9999.0   0   0   #Hnorm
Standard precipitation index:             0     0      0     0    -9999.0   0   0   #SPI
Standard runoff index:                    0     0      0     0    -9999.0   0   0   #SRI
Standardized soil water index:            0     0      0     0    -9999.0   0   0   #SSWI
Standardized ground water index:          0     0      0     0    -9999.0   0   0   #SGWI
Percentile:                               0     0      0     0    -9999.0   0   0   #Percentile
River flow variate:                       0     0      0     0    -9999.0   0   0   #RFV
K-S test:                                 0     0      0     0    -9999.0   0   0   #K-S test
Tendency:                                 0     0      0     0    -9999.0   0   0   #Tendency
Tendency correlation:                     0     0      0     0    -9999.0   0   0   #Tendency correlati
```

- Each line specifies an analysis metric
- 8 entries for each metric

- Metric name (use the user's guide or the master file)
- Use option (0 or 1)
- Time option (0 or 1 ; whether to compute the metric at the temporal averaging interval and output them at the stats output intervals)
- Temporal output - whether to write gridded metric files at the stats output interval (time option must also be enabled)
- Extract time series - whether to extract (ASCII) time series files for the metric, for each sub-domains specified in the time series location file
- Threshold: Threshold value to be used in computing categorical metrics
- Compute average seasonal cycle (monthly, 3-monthly)
- Compute average diurnal cycle

Configuration: Spatial averaging/confidence intervals

```
Spatial averaging mode:                    "pixel-by-pixel"
```

```
Regional mask file for spatial averaging: none
```

```
Confidence interval (%):                          95
```

- "pixel-by-pixel" option computes the metrics separately at each grid point
- "region-based" option computes the metrics using the average values of the datastreams over each region
  - Requires the user to provide a categorical map (in binary, big-endian, sequential access format)

- Confidence interval threshold for computed statistics. The CIs are calculated based on a two-tail t-test
- CIs are computed only across the spatial domain (and not temporally).
  - e.g. if RMSE is computed for 100 stations, then the reported CIs are the values for the average RMSE for the 100 stations.

Configuration: Stratification options

```
Variable-based stratification:                1
Stratification variable:                      'SWdown_f'
Stratification threshold:                     0.1
```

When variable based stratification is used, 3 values will be computed for each metric

1. Metric value with no stratification
2. Metric value where the stratification variable value is above the threshold
3. Metric value where the stratification variable value is below the threshold

```
External data-based stratification:           1
Stratification attributes file:               strat_attribs.txt
```

```
#Number of stratification data sources
3
#Stratification data files
srtm_elev1km.1gd4r
srtm_slope1km.1gd4r
srtm_aspect1km.1gd4r
#stratifcation variable name
ELEV
SLOPE
ASPECT
#Max (row 1) min (row2) values for each category
7000 1.0 6
500  0.0 0
#number of bins
12 12 12
```

Stratification performed for three data sources

Separate files (for each metric) that computes metric values for the specified number of bins will be generated.

e.g. RMSE for each of the 12 elevation, slope and aspect categories will be computed

Configuration: Smoothing

```
Apply temporal smoothing to obs:                    1
Obs temporal smoothing window half length:          "2da"
Obs temporal smoothing window interval:             "1da"
```

`Obs temporal smoothing window half length:` specifies the observation temporal smoothing window half length. The smoothing window is then defined as (current time +/- half length).

`Obs temporal smoothing window interval:` specifies the observation temporal smoothing window interval. This will be used as the increment length across the smoothing window. For e.g., if the window half length is specified as 2 days the smoothing window will be of 5 days. If the smoothing window interval is 1 day, then number of points in the smoothing window will be 5 (-2 da, -1da, current day, +1da, +2da).

Finally, details of the datastreams…

```
LIS output interval:                              "1da"
LIS output analysis data class:                   "LSM"
LIS output number of surface model types:          1
LIS output surface model types:                   "LSM"
LIS output model name:                            "Noah.3.3"
LIS output domain and parameter file:           ../DATA_Noah33_CONUS/lis_input.d01.nc
LIS output directory:                           ../DATA_Noah33_CONUS/OUTPUT
LIS output naming style:                        "3 level hierarchy"
LIS output methodology:                         "2d gridspace"
LIS output format:                              "netcdf"
LIS output attributes file:                     ../DATA_Noah33_CONUS/NOAH33_OUTPUT_LIST.TBL
LIS output maximum number of surface type tiles per grid:     1
LIS output minimum cutoff percentage (surface type tiles):    0.10
LIS output maximum number of soil texture tiles per grid:     1
LIS output minimum cutoff percentage (soil texture tiles):    0.10
LIS output maximum number of soil fraction tiles per grid:    1
LIS output minimum cutoff percentage (soil fraction tiles):   0.10
LIS output maximum number of elevation bands per grid:        1
LIS output minimum cutoff percentage (elevation bands):       0.10
LIS output maximum number of slope bands per grid:            1
LIS output minimum cutoff percentage (slope bands):           0.10
LIS output maximum number of aspect bands per grid:           1
LIS output minimum cutoff percentage (aspect bands):          0.10
LIS output number of ensembles per tile:                  1
LIS output nest index:                                    1
LIS output elevation data source:                        none
LIS output slope data source:                            none
LIS output aspect data source:                           none
LIS output soil texture data source:                     none
LIS output soil fraction data source:                    none
LIS output number of soil moisture layers:              4
LIS output number of soil temperature layers:           4
LIS output soil moisture layer thickness:             0.1 0.3 0.6 1.0
LIS output soil temperature layer thickness:          0.1 0.3 0.6 1.0
```

```
ARS soil moisture observation directory: ../DATA_ARS_Watersheds/
ARS soil moisture station list file:      ../DATA_ARS_Watersheds/stnlist.dat
```

Description of the LIS output

- Output interval
- Analysis data class (LSM, routing, RTM, Irrigation, ..)
- LDT generated input file used in the LIS run
- Output attributes table used in the LIS run
- Output naming style, format, methodology
- Subgrid tiling settings used in the LIS run
- Soil layering information

Description of the ARS data

- Data directory, list of stations

Running LVT (Example 1)

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 %
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % ls
lis_input.nldas.nc*  LVT*  lvt.config*  lvtlog.0000*  METRICS.TBL*  STATS/  TS_LOCATIONS.TXT*
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 %
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % tail lvtlog.0000
 [INFO] LVT cycle time: 08/30/2006 23:30:00
 [ERR] LIS file
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200608/LIS_HIST_200608302330.d01.nc
  does not exist
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 08/31/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200608/LIS_HIST_200608310000.d01.nc
 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 %
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % ls STATS/
ACORR_SUMMARY_STATS.dat              MEAN_lr.dat   MEAN_SUMMARY_STATS.dat
LVT_ACORR_FINAL.200608310000.d01.nc  MEAN_lw.dat   MEAN_wg.dat
LVT_MEAN_FINAL.200608310000.d01.nc   MEAN_rc.dat   RST/
```

Run the executable!

The run generates a LVT logfile, STATS output directory

Check to see if the simulation exited cleanly

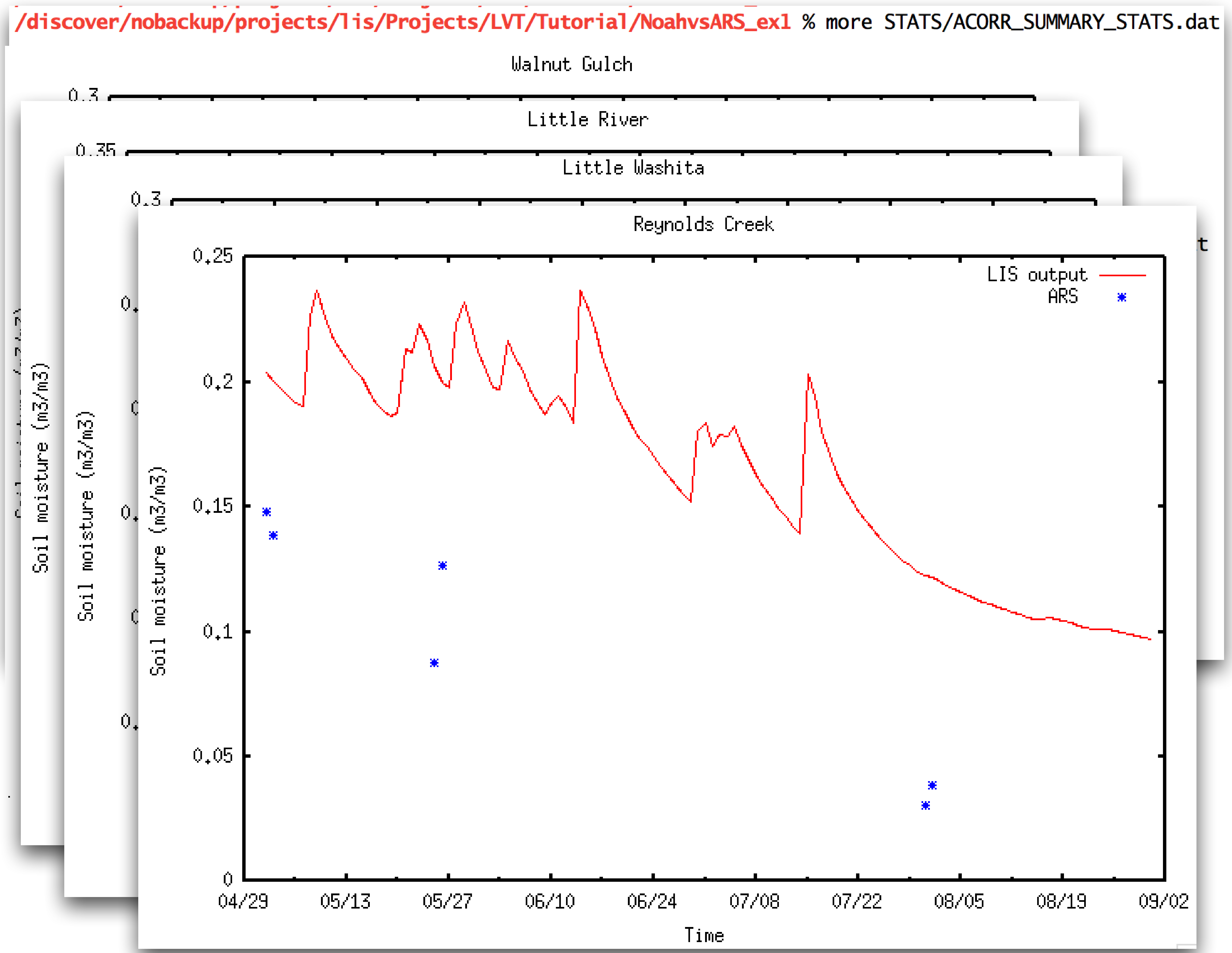
<MEAN/ACORR>_SUMMARY_STATS.dat -files containing summary statistics for mean and anomaly correlation

MEAN_<lr/lw/rc/wg>.dat - files containing mean time series values

LVT_<MEAN/ACORR>_FINAL.200608310000.d01.nc - NetCDF files containing gridded MEAN/Anomaly correlation values for the entire analysis time period

RST - directory containing restart files

Examine the LVT output



- Valid values at Walnut Gulch (wg), Little River (lr), Little Washita (lw)
- Still undefined Anomaly R values at Reynolds Creek (rc)
- Model mean value at rc is valid, observation mean at rc is not
- Likely reason is that the observations at rc are not continuous and therefore do not meet the observation count threshold of 100
- Use the ‘gnuplot’ scripts to plot the timeseries data

  gnuplot wg.plt

  gnuplot lr.plt

  gnuplot lw.plt

  gnuplot rc.plt

ASCII time series file

First 5 columns represent the time information (year, month, day, hour, minute)

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % head STATS/MEAN_wg.dat
2006 05 02 00 00   0.880487E-01 -0.999900E+04  0.880487E-01  0.880487E-01  0.000000E+00 -0.999900E+04
 0.250000E-01  0.700000E-02  0.250000E-01  0.250000E-01 -0.999900E+04 -0.999900E+04
2006 05 03 00 00   0.875005E-01 -0.999900E+04  0.875005E-01  0.875005E-01  0.000000E+00 -0.999900E+04
 0.250000E-01  0.650000E-02  0.250000E-01  0.250000E-01 -0.999900E+04 -0.999900E+04
2006 05 04 00 00   0.869690E-01 -0.999900E+04  0.869690E-01  0.869690E-01  0.000000E+00 -0.999900E+04
 0.250000E-01  0.305556E-02  0.250000E-01  0.250000E-01 -0.999900E+04 -0.999900E+04
2006 05 05 00 00   0.864317E-01 -0.999900E+04  0.864317E-01  0.864317E-01  0.000000E+00 -0.999900E+04
 0.250000E-01  0.111806E-02  0.250000E-01  0.250000E-01 -0.999900E+04 -0.999900E+04
```

For each variable, 6 columns for datastream1, 6 columns for data stream 2

Columns 6-11 represent the soil moisture values from LIS output, 12-17 represent soil moisture values from ARS data

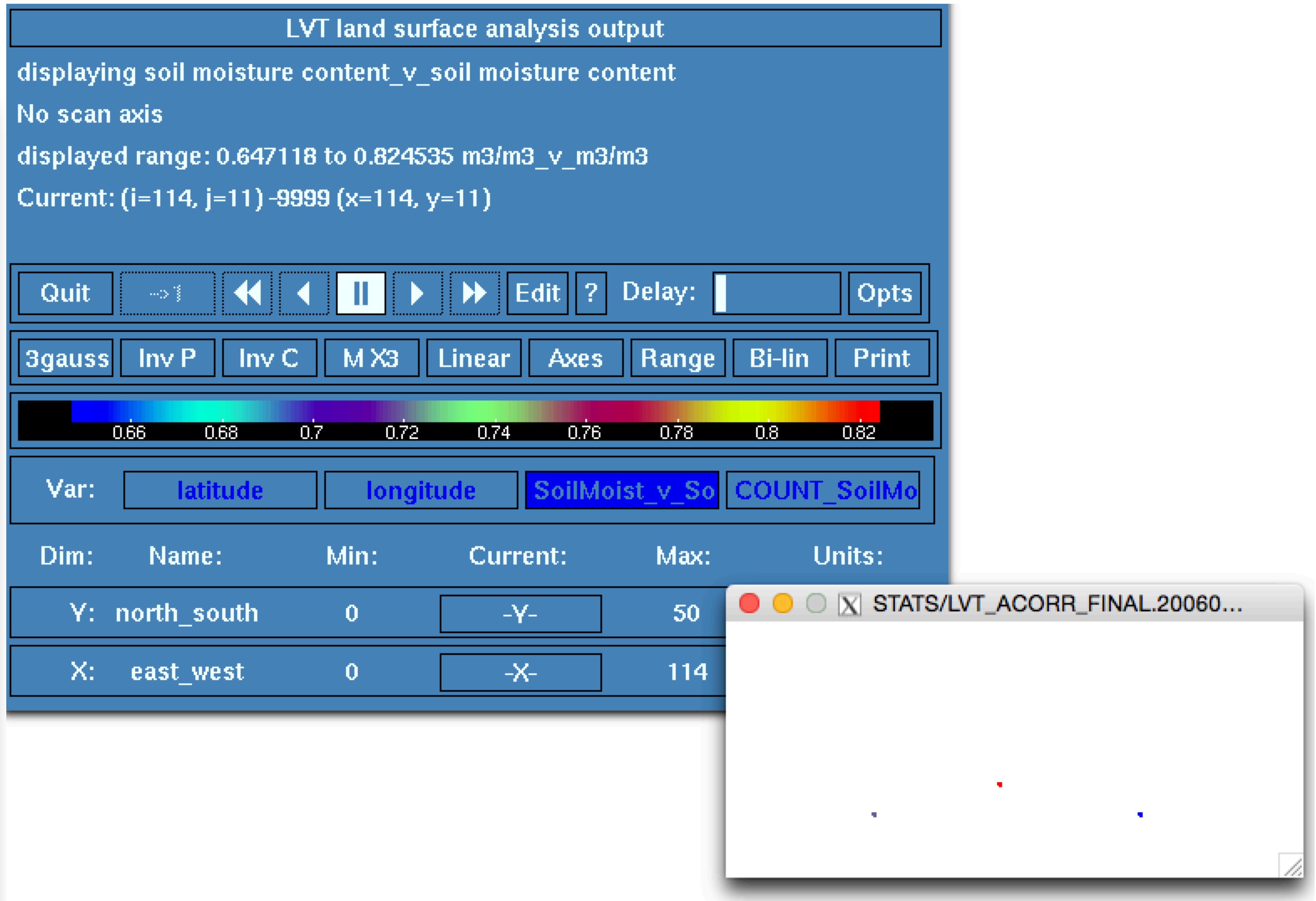
- ☑ mean value
- ☑ standard deviation
- ☑ minimum
- ☑ maximum
- ☑ ensemble standard deviation
- ☑ confidence interval

If more variables are included in the analysis, additional columns will be included for each variable

Note that for comparison metrics, there will be no columns for observation values

FINAL NetCDF files

```
netcdf LVT_ACORR_FINAL.200608310000.d01 {
dimensions:
        east_west = 115 ;
        north_south = 51 ;
        time = 1 ;
variables:
        float latitude(north_south, east_west) ;
                latitude:units = "degree_north" ;
                latitude:standard_name = "latitude" ;
                latitude:long_name = "latitude" ;
                latitude:scale_factor = 1.f ;
                latitude:add_offset = 0.f ;
                latitude:missing_value = -9999.f ;
                latitude:_FillValue = -9999.f ;
        float longitude(north_south, east_west) ;
                longitude:units = "degree_east" ;
                longitude:standard_name = "longitude" ;
                longitude:long_name = "longitude" ;
                longitude:scale_factor = 1.f ;
                longitude:add_offset = 0.f ;
                longitude:missing_value = -9999.f ;
                longitude:_FillValue = -9999.f ;
        float time(time) ;
                time:units = "minutes since 2006-08-31 00:00:00" ;
                time:long_name = "time" ;
                time:time_increment = "1800" ;
                time:begin_date = "20060831" ;
                time:begin_time = "000000" ;
        float SoilMoist_v_SoilMoist(north_south, east_west) ;
                SoilMoist_v_SoilMoist:units = "m3/m3_v_m3/m3" ;
                SoilMoist_v_SoilMoist:standard_name = "soil_moisture_content_v_soil_moisture_content" ;
                SoilMoist_v_SoilMoist:long_name = "soil moisture content_v_soil moisture content" ;
                SoilMoist_v_SoilMoist:scale_factor = 1.f ;
                SoilMoist_v_SoilMoist:add_offset = 0.f ;
                SoilMoist_v_SoilMoist:missing_value = -9999.f ;
                SoilMoist_v_SoilMoist:_FillValue = -9999.f ;
        float COUNT_SoilMoist_v_SoilMoist(north_south, east_west) ;
                COUNT_SoilMoist_v_SoilMoist:units = "-" ;
                COUNT_SoilMoist_v_SoilMoist:standard_name = "COUNT_soil_moisture_content_v_soil_moisture_content" ;
                COUNT_SoilMoist_v_SoilMoist:long_name = "Number of points in soil moisture content_v_soil moisture content" ;
                COUNT_SoilMoist_v_SoilMoist:scale_factor = 1.f ;
                COUNT_SoilMoist_v_SoilMoist:add_offset = 0.f ;
                COUNT_SoilMoist_v_SoilMoist:missing_value = -9999.f ;
                COUNT_SoilMoist_v_SoilMoist:_FillValue = -9999.f ;
```

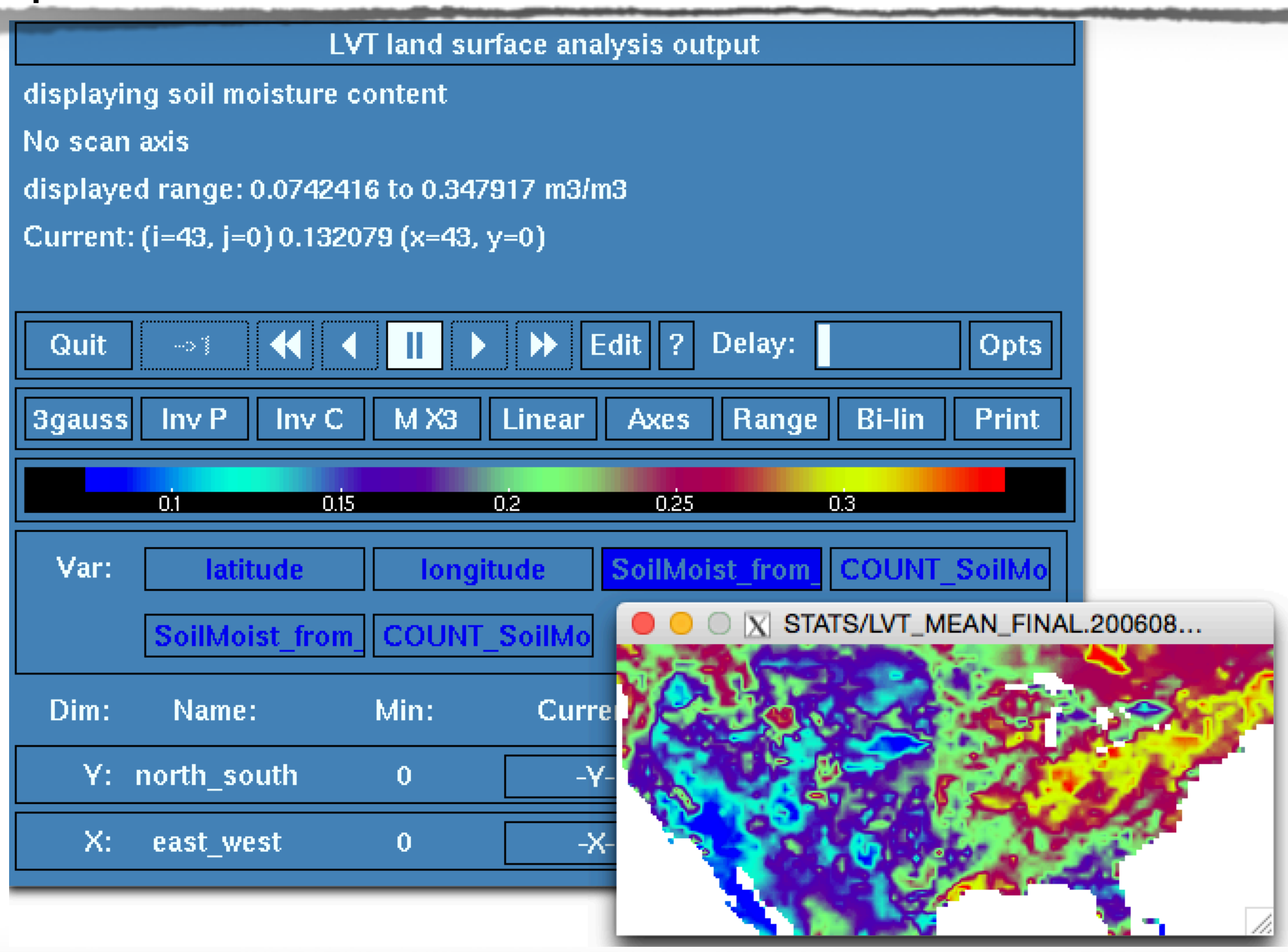


The Anomaly R file contains metric values for three grid points

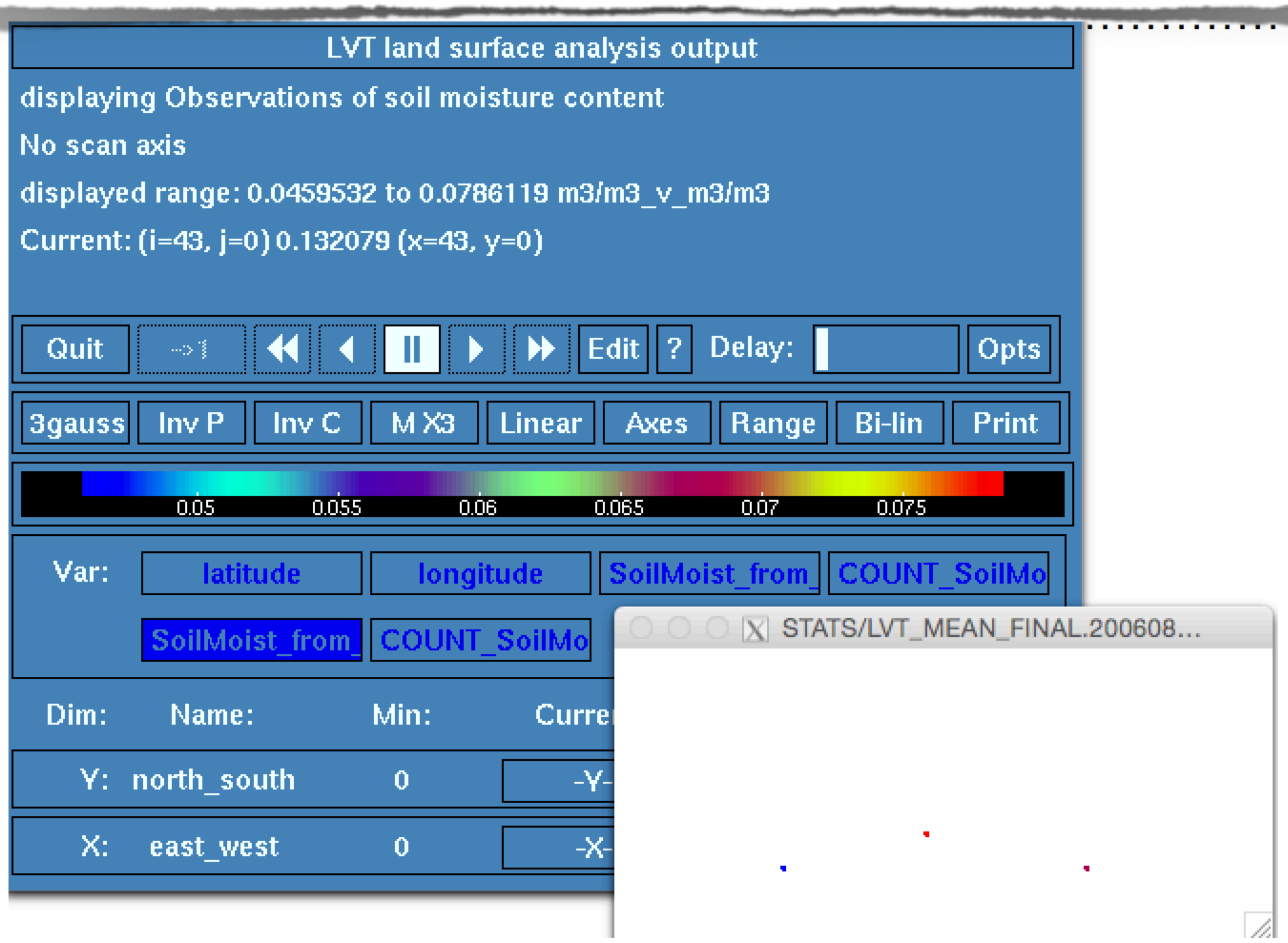
FINAL NetCDF files

```
netcdf LVT_MEAN_FINAL.200608310000.d01 {
dimensions:
        east_west = 115 ;
        north_south = 51 ;
        time = 1 ;
variables:
        float latitude(north_south, east_west) ;
                latitude:units = "degree_north" ;
                latitude:standard_name = "latitude" ;

    float SoilMoist_from_SoilMoist_v_SoilMoist_ds1(north_south, east_west) ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:units = "m3/m3" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:standard_name = "soil_moisture_content" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:long_name = "soil moisture content" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:scale_factor = 1.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:add_offset = 0.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:missing_value = -9999.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds1:_FillValue = -9999.f ;
    float COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1(north_south, east_west) ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:units = "-" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:standard_name = "COUNT_soil_moisture_content" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:long_name = "Number of points in soil moisture content" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:scale_factor = 1.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:add_offset = 0.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:missing_value = -9999.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds1:_FillValue = -9999.f ;
    float SoilMoist_from_SoilMoist_v_SoilMoist_ds2(north_south, east_west) ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:units = "m3/m3_v_m3/m3" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:standard_name = "soil_moisture_content" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:long_name = "Observations of soil moisture content" ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:scale_factor = 1.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:add_offset = 0.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:missing_value = -9999.f ;
          SoilMoist_from_SoilMoist_v_SoilMoist_ds2:_FillValue = -9999.f ;
    float COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2(north_south, east_west) ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:units = "-" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:standard_name = "COUNT_soil_moisture_content" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:long_name = "Number of observation points of soil moisture content" ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:scale_factor = 1.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:add_offset = 0.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:missing_value = -9999.f ;
          COUNT_SoilMoist_from_SoilMoist_v_SoilMoist_ds2:_FillValue = -9999.f ;
```

Soil moisture field from ds1 (datastream1) in the comparison of soil moisture vs. soil moisture



Soil moisture field from ds2 (datastream2) in the comparison of soil moisture vs. soil moisture



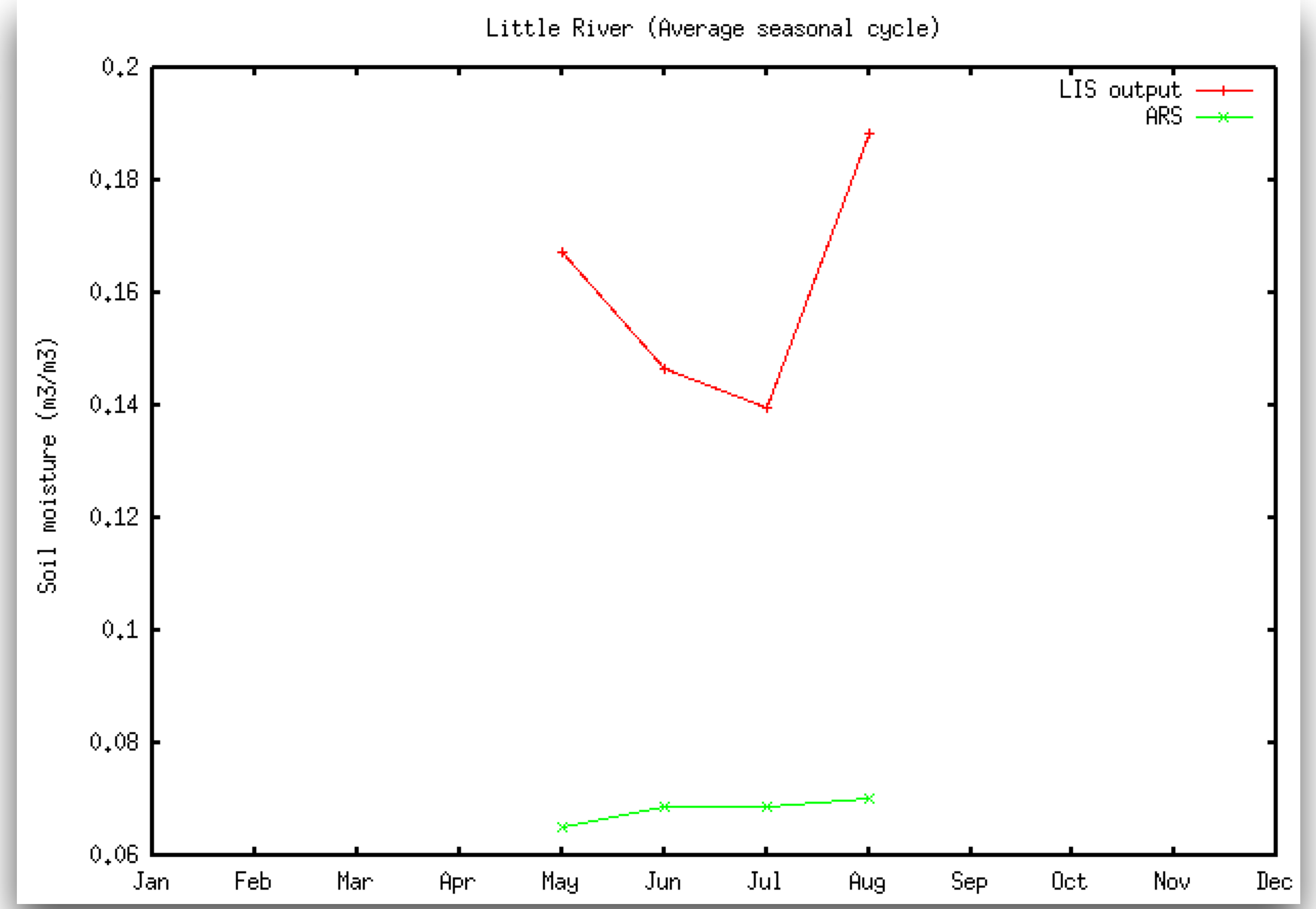
Lets look at the average seasonal cycle..

```
#name                total in-time writeTS extractTS threshold SC ADC short_name
Mean:                                   1    1     0     1    -9999.0  1   0   #Mean
Anomaly:                                0    0     0     0    -9999.0  0   0   #Anomaly
Standard deviation:                     0    0     0     0    -9999.0  0   0   #Std
```

Enables the computation of average seasonal cycles

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NoahvsARS_ex1 % ls STATS/*ASC*
STATS/MEAN_ASC_lr_SoilMoist_v_SoilMoist.dat   STATS/MEAN_ASC_rc_SoilMoist_v_SoilMoist.dat
STATS/MEAN_ASC_lw_SoilMoist_v_SoilMoist.dat   STATS/MEAN_ASC_wg_SoilMoist_v_SoilMoist.dat
```

gnuplot lr_asc.plt



Example 2

Comparing two LIS outputs:
LIS Noah LSM output vs. LIS CLSM LSM output

Example 2: Comparing two LIS outputs..

Noah LSM output (daily) vs CLSM LSM output (3 hourly)

```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from a
# LIS output (from Noah.3.3 LSM) against variables from another LIS output
# (CLSM LSM)
#
# The model output from the Noah LSM is produced over the CONUS domain at
# 0.125 deg spatial resolution (at a daily interval) whereas the CLSM
# output is produced globally at 0.25 deg spatial resolution (at 3hr
# intervals). The LVT analysis is conducted over a CONUS domain at 0.5 deg
# spatial resolution.
#
# The following variables are compared: Qle, Qh, SoilMoist (2 layers) and
# rootzone soil moisture
#
# The following metrics are used: RMSE, Bias
#
#------------------------------------------------------------------------------

LVT running mode:                    "Data intercomparison"
Map projection of the LVT analysis: "latlon"
Analysis data class:                "LSM"
LVT output format:                  "netcdf"
LVT output methodology:             "2d gridspace"
Analysis data sources:             "LIS output" "LIS output"
```

Both data sources are “LIS output”

Datastream specification (when the two analysis sources are the same)

```
LIS output number of surface model types:      1   1
LIS output surface model types:               "LSM" "LSM"
LIS output analysis data class:               "LSM" "LSM"
LIS output model name:                        "Noah.3.3" "CLSM F2.5"
LIS output map projection:                    "latlon" "latlon"
LIS output domain and parameter file:        ../DATA_Noah33_CONUS/lis_input.d01.nc ../DATA_CLSM_GLB/lis_input.d01.nc
LIS output directory:                        ../DATA_Noah33_CONUS/OUTPUT ../DATA_CLSM_GLB/OUTPUT
LIS output naming style:              "3 level hierarchy" "3 level hierarchy"
LIS output methodology:               "2d gridspace"  "2d gridspace"
LIS output format:                    "netcdf"  "netcdf"
LIS output attributes file:    ../DATA_Noah33_CONUS/NOAH33_OUTPUT_LIST.TBL ../DATA_CLSM_GLB/CLSM_OUTPUT_LIST.TBL
LIS output maximum number of surface type tiles per grid:     1     1
LIS output minimum cutoff percentage (surface type tiles):    0.10  0.10
LIS output maximum number of soil texture tiles per grid:     1     1
LIS output minimum cutoff percentage (soil texture tiles):    0.10  0.10
LIS output maximum number of soil fraction tiles per grid:    1     1
LIS output minimum cutoff percentage (soil fraction tiles):   0.10  0.10
LIS output maximum number of elevation bands per grid:        1     1
LIS output minimum cutoff percentage (elevation bands):       0.10  0.10
LIS output maximum number of slope bands per grid:            1     1
LIS output minimum cutoff percentage (slope bands):           0.10  0.10
LIS output maximum number of aspect bands per grid:           1     1
LIS output minimum cutoff percentage (aspect bands):          0.10  0.10
LIS output number of ensembles per tile:  1 1
LIS output nest index:      1   1
LIS output elevation data source: none none
LIS output slope data source:     none none
LIS output aspect data source:     none none
LIS output soil texture data source: none none
LIS output soil fraction data source: none none
LIS output number of soil moisture layers:  4 3
LIS output number of soil temperature layers:  4 6
LIS output soil moisture layer thickness: 0.1 0.3 0.6 1.0    0.02 1.0 2.0
LIS output soil temperature layer thickness:  0.1 0.3 0.6 1.0   0.05 0.10 0.19 0.39 0.76 1.51
```

The analysis sources are specified in columns, separated by spaces

Noah has 4 soil moisture layers, CLSM has 3

Soil moisture layer thickness of Noah (4 values) are listed first followed by that of CLSM (3 values)

Similar specification for the soil temperature layers (4 & 6 values)

Examining example 2 output..

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/CLSMvsNoah_ex2 % more STATS/RMSE_SUMMARY_STATS.dat
 ----------------------------------------------------------
 VAR: surface_upward_latent_heat_flux_v_surface_upward_latent_heat_flux
 ----------------------------------------------------------
      ALL:     0.343E+02 +/-     0.531E+00        4435
HIGHPLAINS     0.305E+02 +/-     0.113E+01         434
WESTCOAST:     0.423E+02 +/-     0.258E+01         273
       SGP:    0.265E+02 +/-     0.176E+01         325
  MIDWEST:     0.370E+02 +/-     0.102E+01         476
NORTHEAST:     0.526E+02 +/-     0.195E+01         236
 ----------------------------------------------------------
 VAR: surface_upward_sensible_heat_flux_v_surface_upward_sensible_heat_flux
 ----------------------------------------------------------
      ALL:     0.377E+02 +/-     0.555E+00        4435
HIGHPLAINS     0.337E+02 +/-     0.169E+01         434
WESTCOAST:     0.497E+02 +/-     0.262E+01         273
       SGP:    0.304E+02 +/-     0.152E+01         325
  MIDWEST:     0.359E+02 +/-     0.117E+01         476
NORTHEAST:     0.437E+02 +/-     0.212E+01         236
 ----------------------------------------------------------
 VAR: soil_moisture_content_v_soil_moisture_content
 ----------------------------------------------------------
      ALL:     0.459E-01 +/-     0.805E-03        4435
HIGHPLAINS     0.392E-01 +/-     0.213E-02         434
WESTCOAST:     0.591E-01 +/-     0.477E-02         273
       SGP:    0.434E-01 +/-     0.313E-02         325
  MIDWEST:     0.393E-01 +/-     0.191E-02         476
NORTHEAST:     0.454E-01 +/-     0.360E-02         236
 ----------------------------------------------------------
 VAR: root_zone_soil_moisture_v_root_zone_soil_moisture
 ----------------------------------------------------------
      ALL:     0.430E-01 +/-     0.868E-03        4435
HIGHPLAINS     0.443E-01 +/-     0.284E-02         434
WESTCOAST:     0.501E-01 +/-     0.405E-02         273
       SGP:    0.435E-01 +/-     0.319E-02         325
  MIDWEST:     0.425E-01 +/-     0.264E-02         476
NORTHEAST:     0.490E-01 +/-     0.411E-02         236
```

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/CLSMvsNoah_ex2 % more STATS/BIAS_SUMMARY_STATS.dat
 ----------------------------------------------------------
 VAR: surface_upward_latent_heat_flux_v_surface_upward_latent_heat_flux
 ----------------------------------------------------------
      ALL:    -0.123E+02 +/-     0.654E+00        4435
HIGHPLAINS    -0.130E+02 +/-     0.139E+01         434
WESTCOAST:    -0.251E+02 +/-     0.294E+01         273
       SGP:   -0.161E+02 +/-     0.208E+01         325
  MIDWEST:    -0.164E+02 +/-     0.128E+01         476
NORTHEAST:    -0.316E+02 +/-     0.231E+01         236
 ----------------------------------------------------------
 VAR: surface_upward_sensible_heat_flux_v_surface_upward_sensible_heat_flux
 ----------------------------------------------------------
      ALL:     0.118E+02 +/-     0.834E+00        4435
HIGHPLAINS     0.483E+01 +/-     0.241E+01         434
WESTCOAST:     0.345E+02 +/-     0.331E+01         273
       SGP:    0.310E+01 +/-     0.238E+01         325
  MIDWEST:     0.110E+02 +/-     0.181E+01         476
NORTHEAST:     0.255E+02 +/-     0.258E+01         236
 ----------------------------------------------------------
 VAR: soil_moisture_content_v_soil_moisture_content
 ----------------------------------------------------------
      ALL:    -0.185E-01 +/-     0.133E-02        4435
HIGHPLAINS    -0.163E-01 +/-     0.345E-02         434
WESTCOAST:    -0.448E-01 +/-     0.626E-02         273
       SGP:   -0.271E-01 +/-     0.443E-02         325
  MIDWEST:    -0.157E-01 +/-     0.319E-02         476
NORTHEAST:    -0.149E-01 +/-     0.609E-02         236
 ----------------------------------------------------------
 VAR: root_zone_soil_moisture_v_root_zone_soil_moisture
 ----------------------------------------------------------
      ALL:    -0.135E-01 +/-     0.146E-02        4435
HIGHPLAINS    -0.168E-01 +/-     0.473E-02         434
WESTCOAST:    -0.206E-01 +/-     0.670E-02         273
       SGP:   -0.325E-01 +/-     0.445E-02         325
  MIDWEST:    -0.250E-02 +/-     0.459E-02         476
NORTHEAST:     0.125E-01 +/-     0.722E-02         236
```

Note the order of variables in the output : Qle_vs_Qle, Qh_vs_Qh, soilmoist_vs_soilmoist, rootmoist_vs_rootmoist

The same order will be maintained in the time series and gridded output files

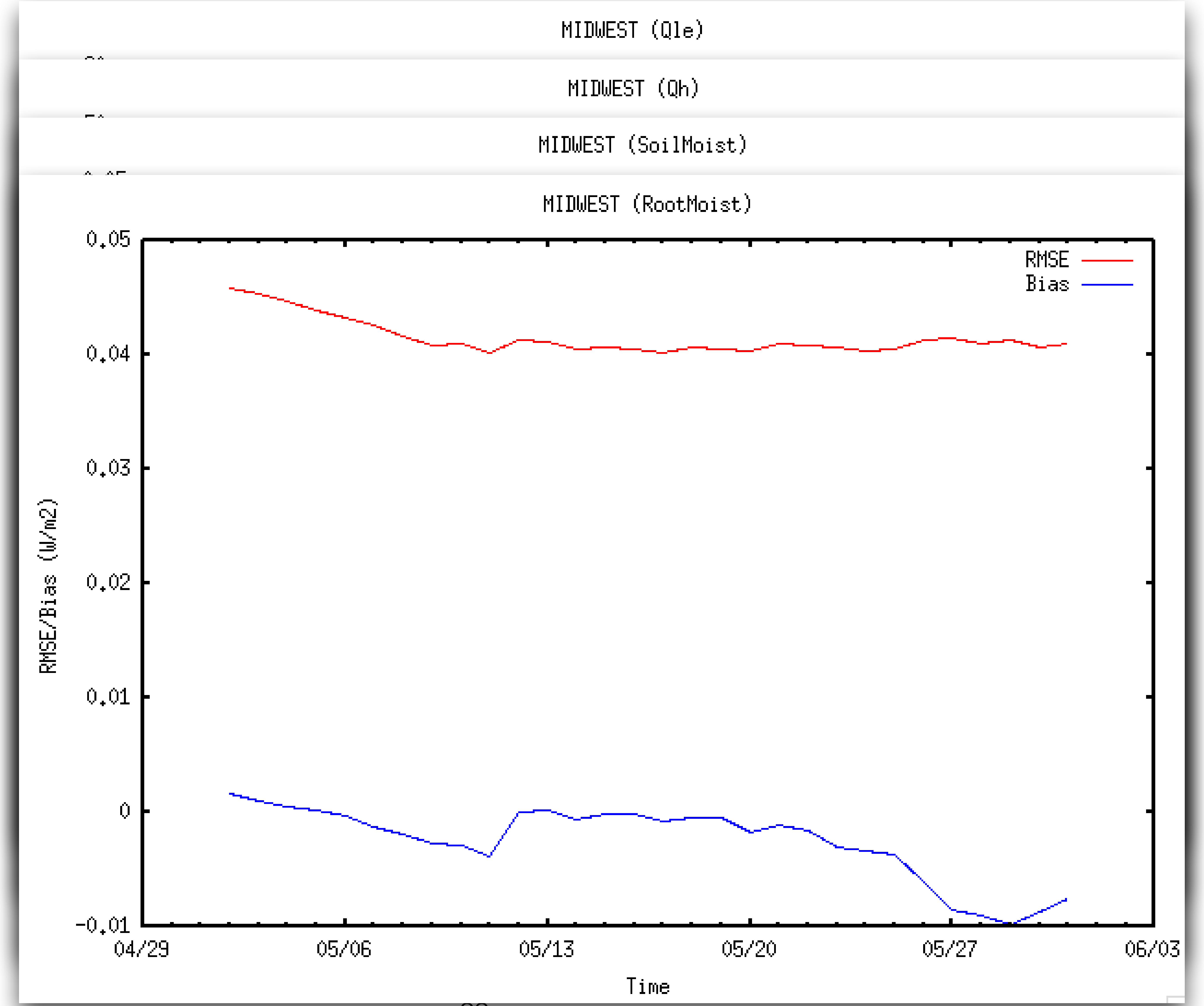
Example 2 time series output..

gnuplot midwest_qle.plt

gnuplot midwest_qh.plt

gnuplot midwest_soilmoist.plt

gnuplot midwest_rootmoist.plt



Example 2 time series output..

```
/discover/[illegible]/projects/lis/Projects/LVT/Tutorial/CLSMvsNoah_ex2 % more STATS/RMSE_MIDWEST.dat
 2006 05 02 00 00   0.303073E+02  0.307827E+02  0.114544E+00  0.136232E+03  0.000000E+00  0.277538E+01  0.329897E+02  0.300064E+02  0.339737E-01  0.162460E+03  0.000000E+00  0.270539E+01  0.343947E-01  0.288456E-01  0.140369E-04  0.160651E+00  0.000000E+00
  0.260073E-02  0.458427E-01  0.344670E-01  0.146329E-04  0.148757E+00  0.000000E+00  0.310756E-02
 2006 05 03 00 00   0.242185E+02  0.255548E+02  0.302167E-01  0.135991E+03  0.000000E+00  0.230403E+01  0.243154E+02  0.217613E+02  0.111393E+00  0.128705E+03  0.000000E+00  0.196200E+01  0.340849E-01  0.268565E-01  0.909716E-04  0.150913E+00  0.000000E+00
  0.242139E-02  0.454131E-01  0.337840E-01  0.100940E-03  0.146671E+00  0.000000E+00  0.304598E-02
 2006 05 04 00 00   0.204195E+02  0.156204E+02  0.143452E+00  0.799989E+02  0.000000E+00  0.140835E+01  0.299072E+02  0.220354E+02  0.133583E+00  0.130257E+03  0.000000E+00  0.198672E+01  0.350656E-01  0.276078E-01  0.426173E-05  0.152007E+00  0.000000E+00
  0.248913E-02  0.447887E-01  0.333807E-01  0.789762E-05  0.145720E+00  0.000000E+00  0.300961E-02
 2006 05 05 00 00   0.200995E+02  0.157672E+02  0.894470E-01  0.852467E+02  0.000000E+00  0.142158E+01  0.306736E+02  0.281602E+02  0.427704E-01  0.135208E+03  0.000000E+00  0.253893E+01  0.367597E-01  0.288425E-01  0.131130E-04  0.157837E+00  0.000000E+00
  0.260045E-02  0.439870E-01  0.326826E-01  0.668168E-04  0.143526E+00  0.000000E+00  0.294668E-02
 2006 05 06 00 00   0.245734E+02  0.152202E+02  0.585899E-01  0.878889E+02  0.000000E+00  0.137226E+01  0.220028E+02  0.159310E+02  0.898514E-01  0.891804E+02  0.000000E+00  0.143635E+01  0.373927E-01  0.286851E-01  0.832975E-04  0.151994E+00  0.000000E+00
  0.258626E-02  0.433204E-01  0.321250E-01  0.642538E-04  0.142541E+00  0.000000E+00  0.289640E-02
```

First 5 columns represent the time information (year, month, day, hour, minute)

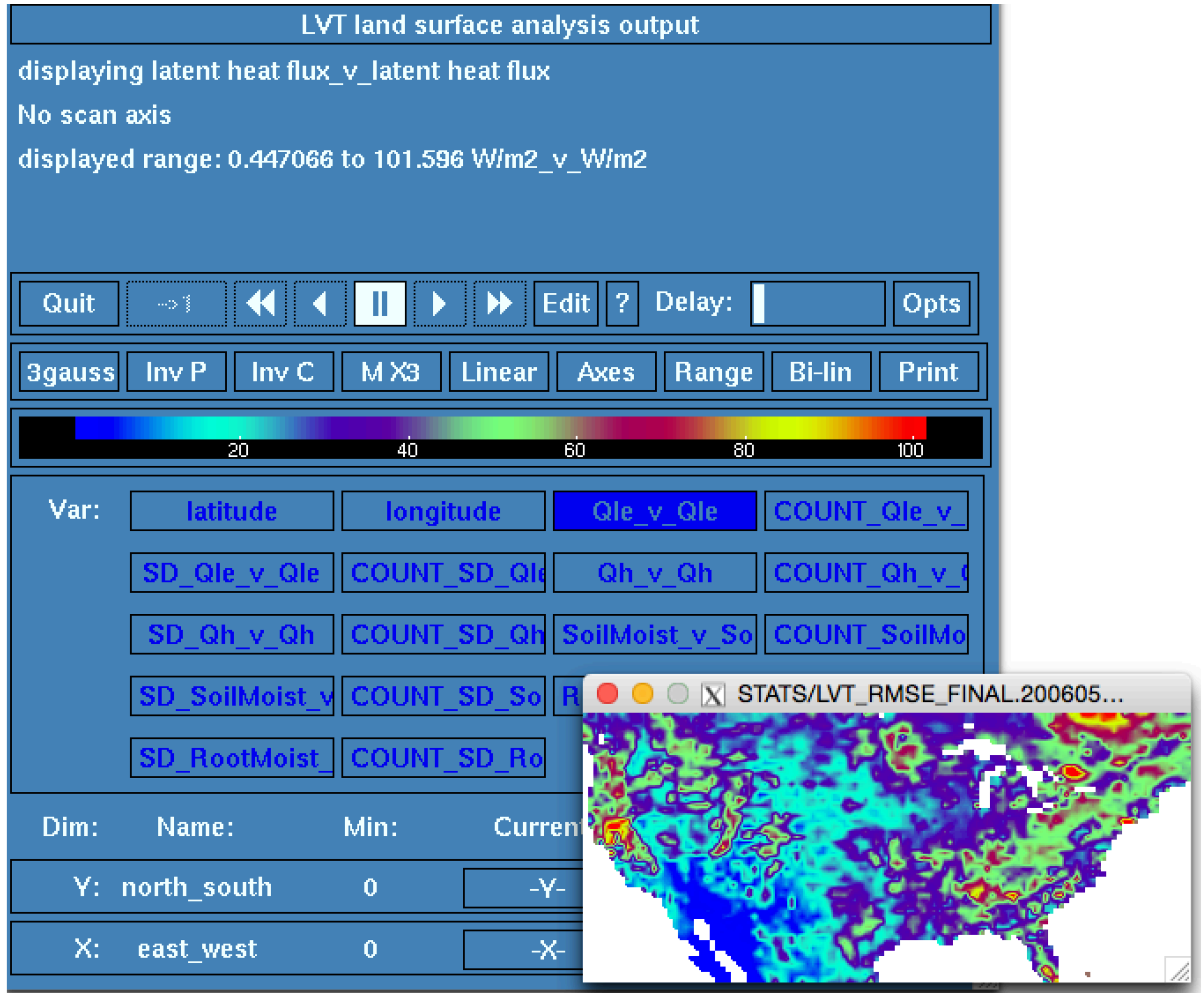
Note that range of standard deviation, min, max, CIs for each variable in each row, as the sub-regions (including MIDWEST) in the TS_LOCATIONS file encompass an area (as opposed to a point in example 1)

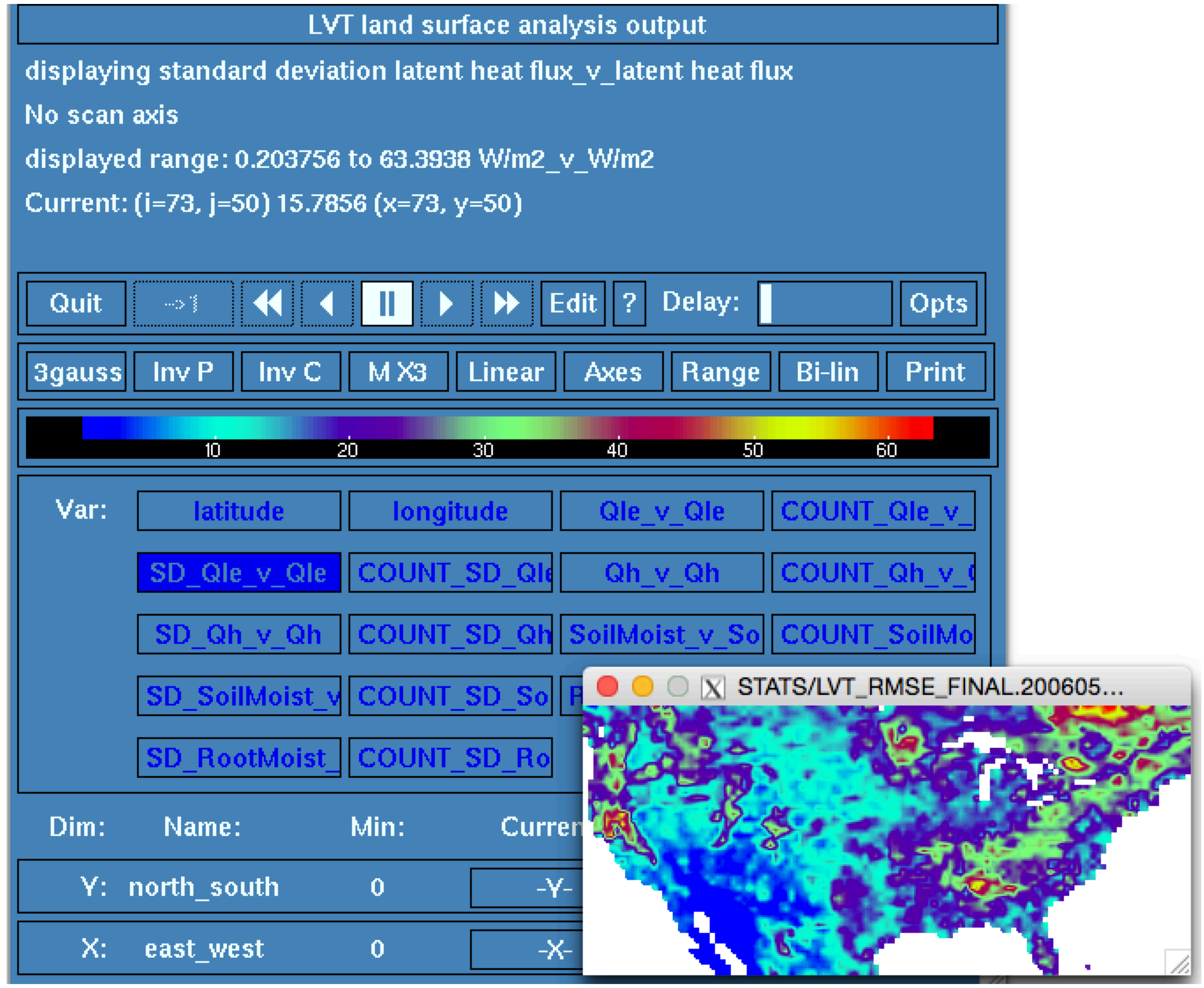
6 columns for RMSE of each variable

Columns 6-11 (Qle), Columns 12-17 (Qh), Columns 18-24 (SoilMoist), Columns 30-35 (RootMoist)

- ☑ mean value
- ☑ standard deviation
- ☑ minimum
- ☑ maximum
- ☑ ensemble standard deviation
- ☑ confidence interval

Example 2 gridded FINAL outputs





For certain metrics (RMSE, for e.g.), the standard deviation of the metric is included in the FINAL output

Example 3

Comparison of two non-LIS outputs: NLDAS2 (Noah) vs. AGRMET operational output

Example 3: NLDA2 vs AGRMET

```
#-----------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from the
# NLDAS2 output against the AGRMET data.
#
# The NLDAS2 data (using Noah LSM) is produced over CONUS at 0.125 deg
# spatial resolution (at hourly intervals). AGRMET outputs are generated
# globally at 3 hourly intervals. Both outputs are generated in grib formats.
# The LVT analysis is conducted
# over a CONUS domain at 0.125 deg spatial resolution.
#
# The following variables are compared: Latent heat flux and surface soil
# moisture
#
# The following metrics are used: Mean and Raw correlation
#
# This configuration includes an example of stratification with an external
# categorical data
#-----------------------------------------------------------------------

LVT running mode:                    "Data intercomparison"
Map projection of the LVT analysis: "latlon"
LVT output format:                 "netcdf"
LVT output methodology:            "2d gridspace"
Analysis data sources:            "NLDAS 2" "AGRMET"
```

NLDAS2 outputs are hourly over CONUS at 0.125 deg, AGRMET outputs are global at 3 hourly intervals at 0.25 deg

Examining example 3 output..

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NLDAS2vsAGRMET_ex3 % more STATS/MEAN_SUMMARY_STATS.dat
 ---------------------------------------------------------
 VAR: surface_upward_latent_heat_flux_v_surface_upward_latent_heat_flux
 ---------------------------------------------------------
     ALL:      0.296E+02 +/-      0.134E+00         75474
HIGHPLAINS     0.262E+02 +/-      0.229E+00          6396
WESTCOAST:     0.378E+02 +/-      0.555E+00          3883
      SGP:     0.242E+02 +/-      0.257E+00          4708
  MIDWEST:     0.268E+02 +/-      0.211E+00          6597
NORTHEAST:     0.245E+02 +/-      0.210E+00          3262
 ---------------------------------------------------------
 VAR: DS2_surface_upward_latent_heat_flux_v_surface_upward_latent_heat_flux
 ---------------------------------------------------------
     ALL:      0.541E+02 +/-      0.213E+00         75591
HIGHPLAINS     0.525E+02 +/-      0.391E+00          6397
WESTCOAST:     0.478E+02 +/-      0.428E+00          3884
      SGP:     0.532E+02 +/-      0.547E+00          4699
  MIDWEST:     0.619E+02 +/-      0.577E+00          6530
NORTHEAST:     0.553E+02 +/-      0.560E+00          3242
 ---------------------------------------------------------
 VAR: soil_moisture_content_v_soil_moisture_content
 ---------------------------------------------------------
     ALL:     -0.568E-01 +/-      0.401E-01         75713
HIGHPLAINS     0.263E+00 +/-      0.308E-01          6397
WESTCOAST:    -0.109E+00 +/-      0.195E+00          3898
      SGP:     0.194E+00 +/-      0.418E-01          4709
  MIDWEST:     0.114E+00 +/-      0.892E-01          6606
NORTHEAST:    -0.229E+00 +/-      0.239E+00          3278
 ---------------------------------------------------------
 VAR: DS2_soil_moisture_content_v_soil_moisture_content
 ---------------------------------------------------------
     ALL:      0.282E+00 +/-      0.703E-03         75591
HIGHPLAINS     0.309E+00 +/-      0.176E-02          6397
WESTCOAST:     0.323E+00 +/-      0.160E-02          3884
      SGP:     0.259E+00 +/-      0.288E-02          4699
  MIDWEST:     0.286E+00 +/-      0.165E-02          6530
NORTHEAST:     0.285E+00 +/-      0.108E-02          3242
```

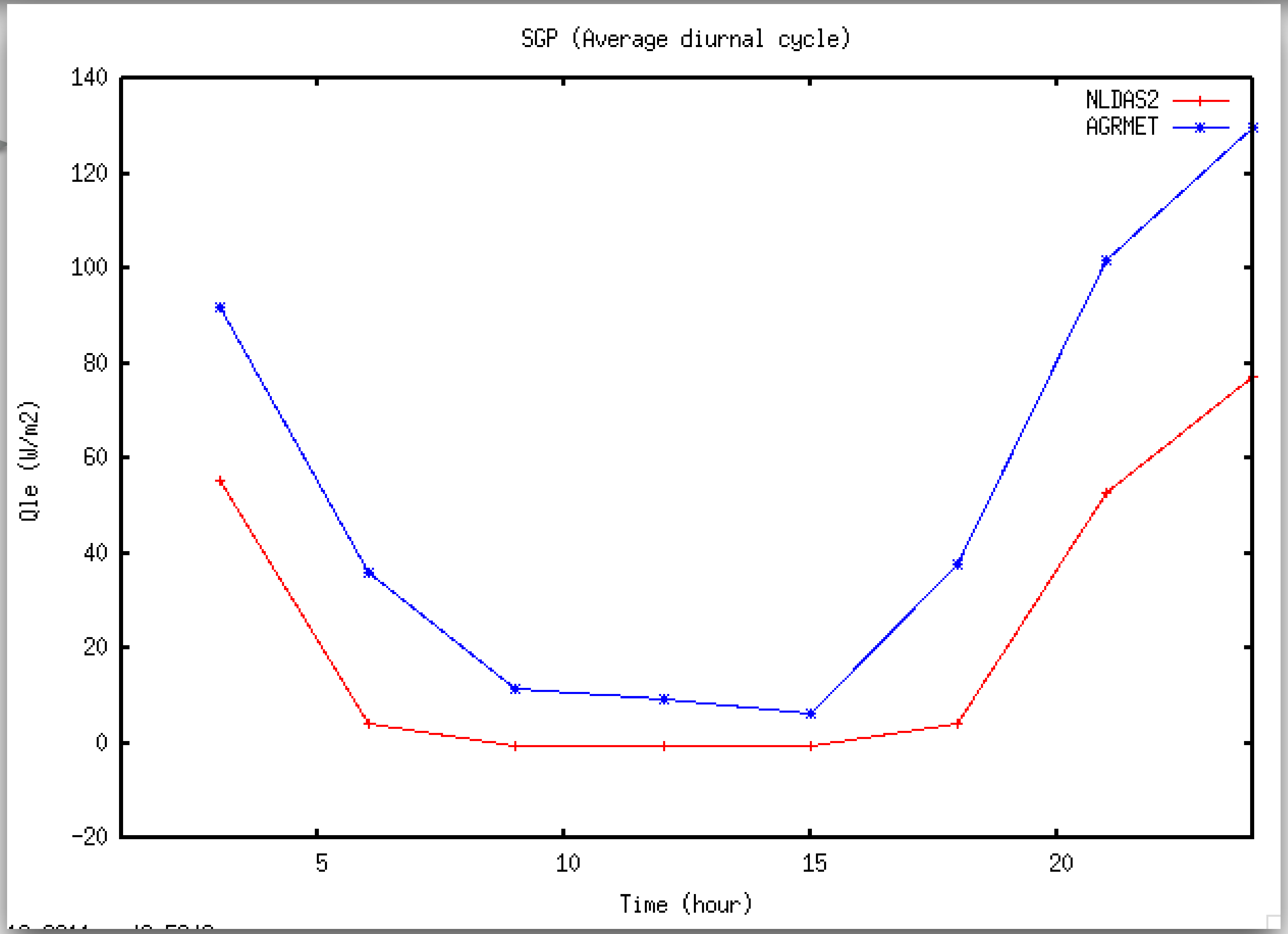
Lets compute the average diurnal cycles..

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NLDAS2vsAGRMET_ex3 % more METRICS.TBL
#name              total in-time writeTS extractTS threshold SC ADC short_name
Mean:                              1    1    0    1   -9999.0  0  1  #Mean
Anomaly:                           0    0    0    0   -9999.0  0  0  #Anomaly
Standard deviation:                0    0    0    0   -9999.0  0  0  #Std
RMSE:                              0    0    0    0   -9999.0  0  0  #RMSE
Bias:                              0    0    0    0   -9999.0  0  0  #Bias
ubRMSE:                            0    0    0    0   -9999.0  0  0  #ubRMSE
Mean absolute error:               0    0    0    0   -9999.0  0  0  #MAE
Anomaly RMSE:                      0    0    0    0   -9999.0  0  0  #ARMSE
```

Enable the computation of average diurnal cycles

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NLDAS2vsAGRMET_ex3 % ls STATS/*ADC*
STATS/MEAN_ADC_HIGHPLAINS_Qle_v_Qle.dat                 STATS/MEAN_ADC_NORTHEAST_Qle_v_Qle.dat                STATS/MEAN_ADC_WESTCOAST_Qle_v_Qle.dat
STATS/MEAN_ADC_HIGHPLAINS_SoilMoist_v_SoilMoist.dat     STATS/MEAN_ADC_NORTHEAST_SoilMoist_v_SoilMoist.dat    STATS/MEAN_ADC_WESTCOAST_SoilMoist_v_SoilMoist.dat
STATS/MEAN_ADC_MIDWEST_Qle_v_Qle.dat                    STATS/MEAN_ADC_SGP_Qle_v_Qle.dat
STATS/MEAN_ADC_MIDWEST_SoilMoist_v_SoilMoist.dat        STATS/MEAN_ADC_SGP_SoilMoist_v_SoilMoist.dat
```

gnuplot sgp_adc.plt



Why are *ADC* files not present for R metric?

Because the temporal computations are not enabled

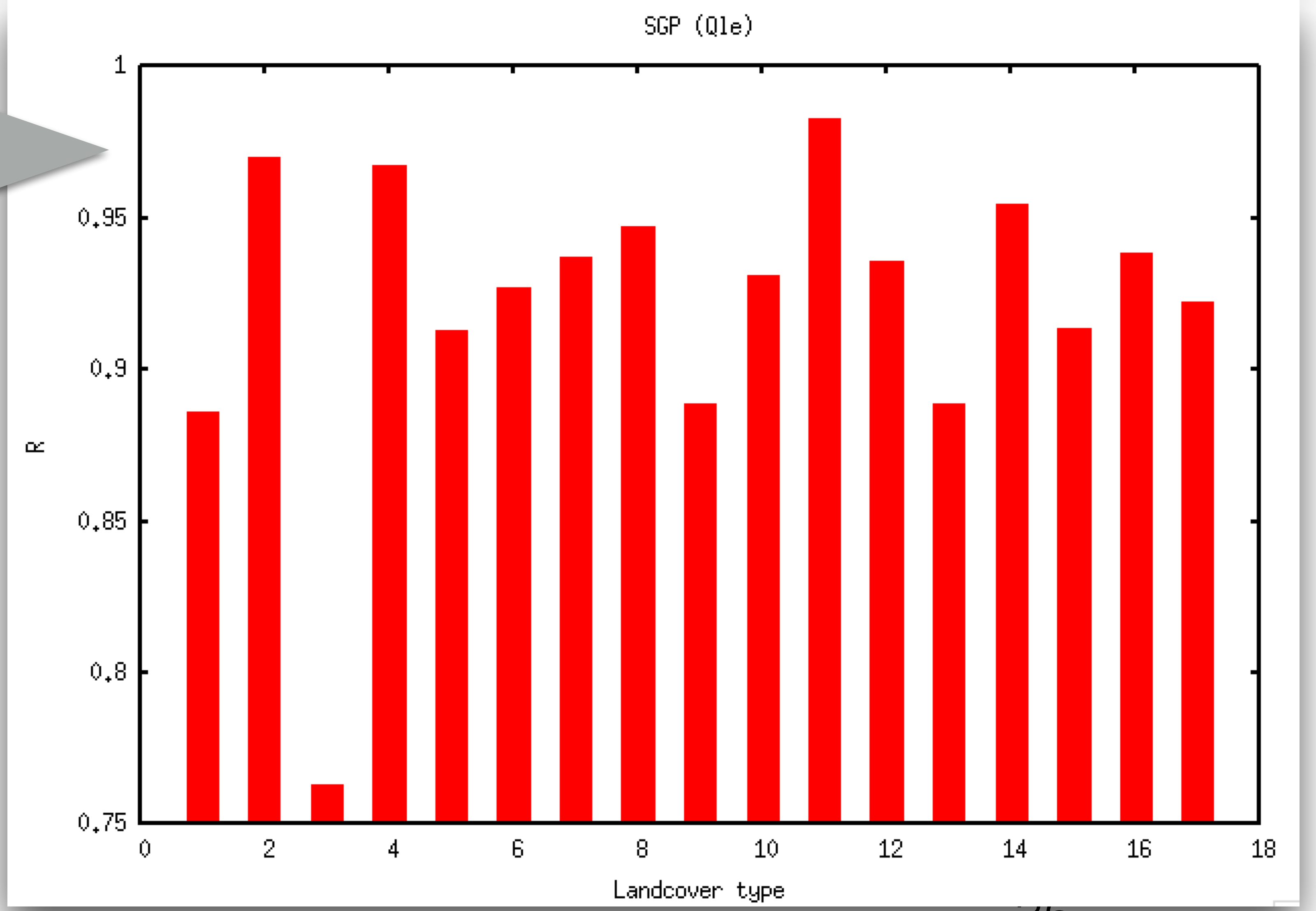
Lets stratify the analysis by landcover type..

```
External data-based stratification:            1
Stratification attributes file:                landcover_strat.txt
```

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NLDAS2vsAGRMET_ex3 % more landcover_strat.txt
#Number of stratification data sources
1
#Stratification data files
landcover_conus.bin
#stratifcation variable name
LANDCOVER
#Max min values
18.0
1.0
#number of bins
17
```

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/NLDAS2vsAGRMET_ex3 % ls STATS/*LANDCOVER*
STATS/MEAN_by_LANDCOVER_Qle_v_Qle.dat               STATS/RCORR_by_LANDCOVER_Qle_v_Qle.dat
STATS/MEAN_by_LANDCOVER_SoilMoist_v_SoilMoist.dat   STATS/RCORR_by_LANDCOVER_SoilMoist_v_SoilMoist.dat
```

gnuplot sgp_r_lc.plt



Example 4

Comparison of two satellite datasets: ESA CCI soil moisture vs GIMMS NDVI

Example 4: ESA CCI vs GIMMS NDVI

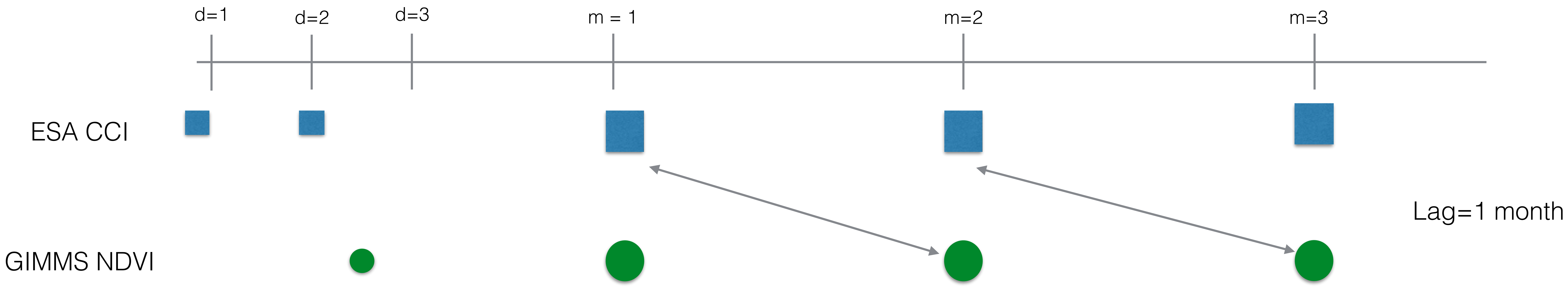
```
#-----------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing the ESA
# CCI soil moisture data against the GIMMS NDVI data
#
# The LVT analysis is conducted over a domain in Africa
# at 0.5 deg spatial resolution.
#
# The following variables are compared: surface soil moisture and NDVI
#
# The following metrics are used: Raw correlation
#
#-----------------------------------------------------------------------------

LVT running mode:                      "Data intercomparison"
Map projection of the LVT analysis: "latlon"
LVT output format:                    "netcdf"
LVT output methodology:               "2d gridspace"
Analysis data sources:              "ESA CCI soil moisture" "GIMMS NDVI"
```

ESA CCI data is daily, GIMMS NDVI data is monthly

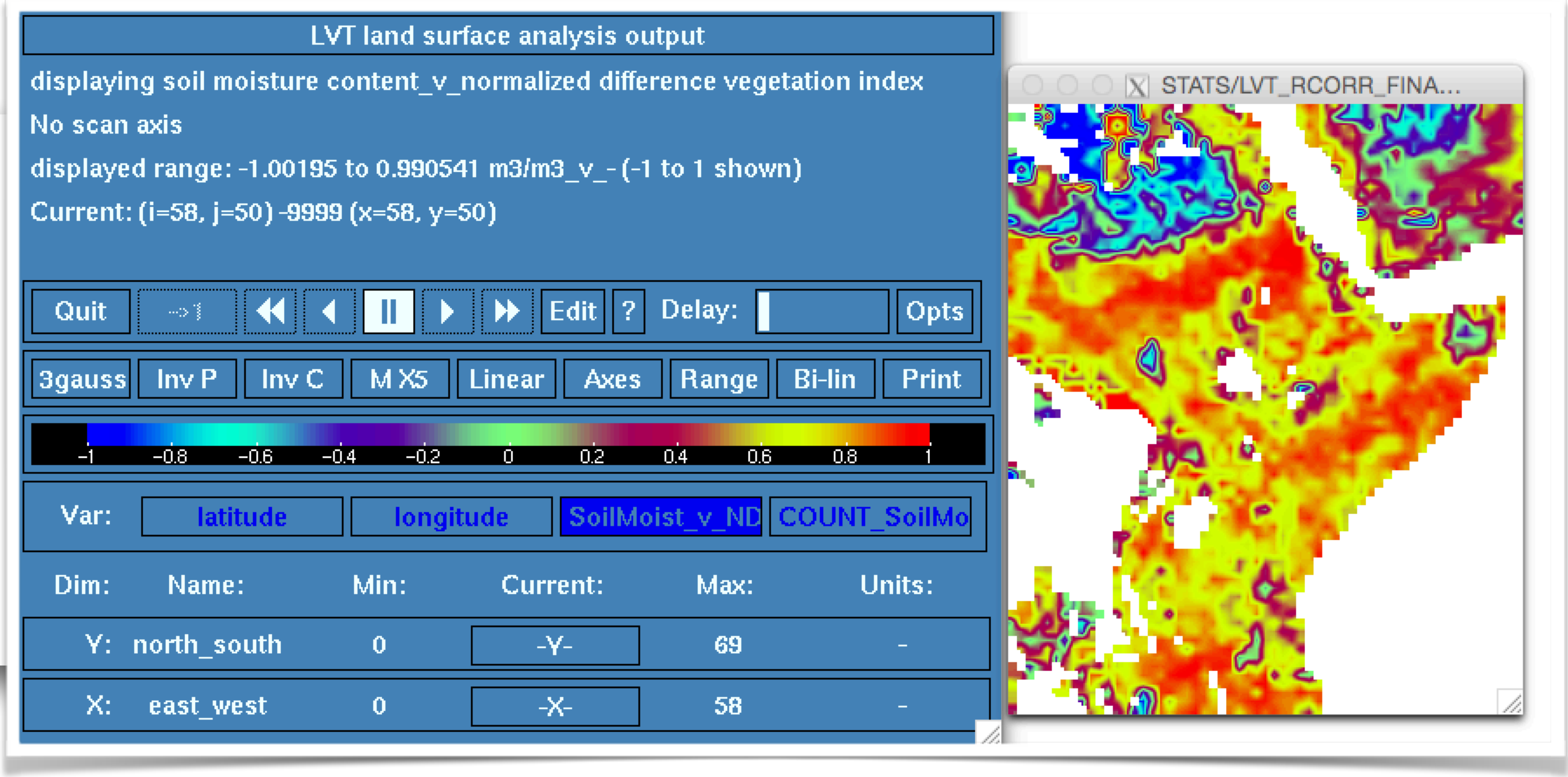
```
Temporal lag in metric computations:              "1mo"
```

The metric (R) is computed with a temporal lag of 1 month - values from datastream 1 will be compared to next month's values from datastream 2

Examining example 4 output..

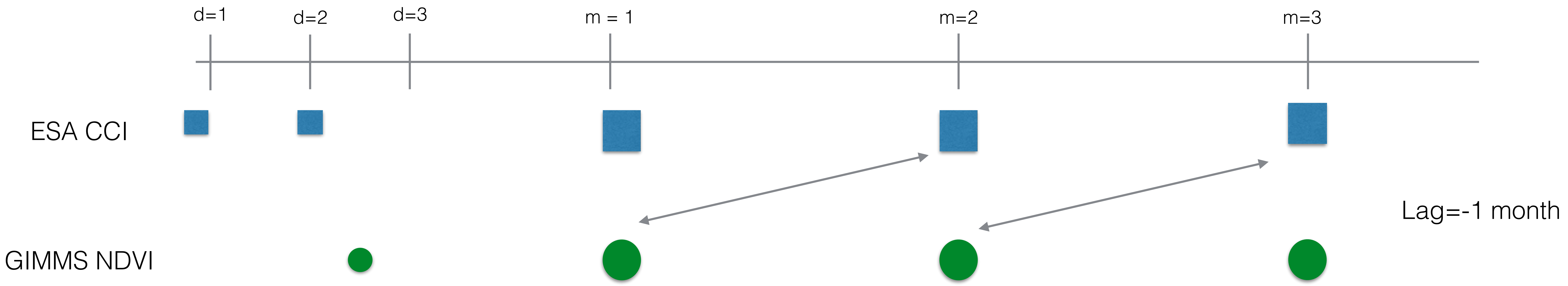



```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % tail lvtlog.0000
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 12/31/2006 00:00:00
 [INFO] Reading ..
 ../DATA_ESACCI/2006/ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED-20061231000000-fv02.0
 .nc
 [INFO] Writing restart file
 ./STATS/RST/LVT.200612310000.rst

 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % more STATS/RCORR_SUMMARY_STATS.dat
 ---------------------------------------------------------
 VAR: soil_moisture_content_v_normalized_difference_vegetation_index
 ---------------------------------------------------------
     ALL:      0.485E+00 +/-     0.184E-01          2659
     REG1:     0.485E+00 +/-     0.184E-01          2659
```

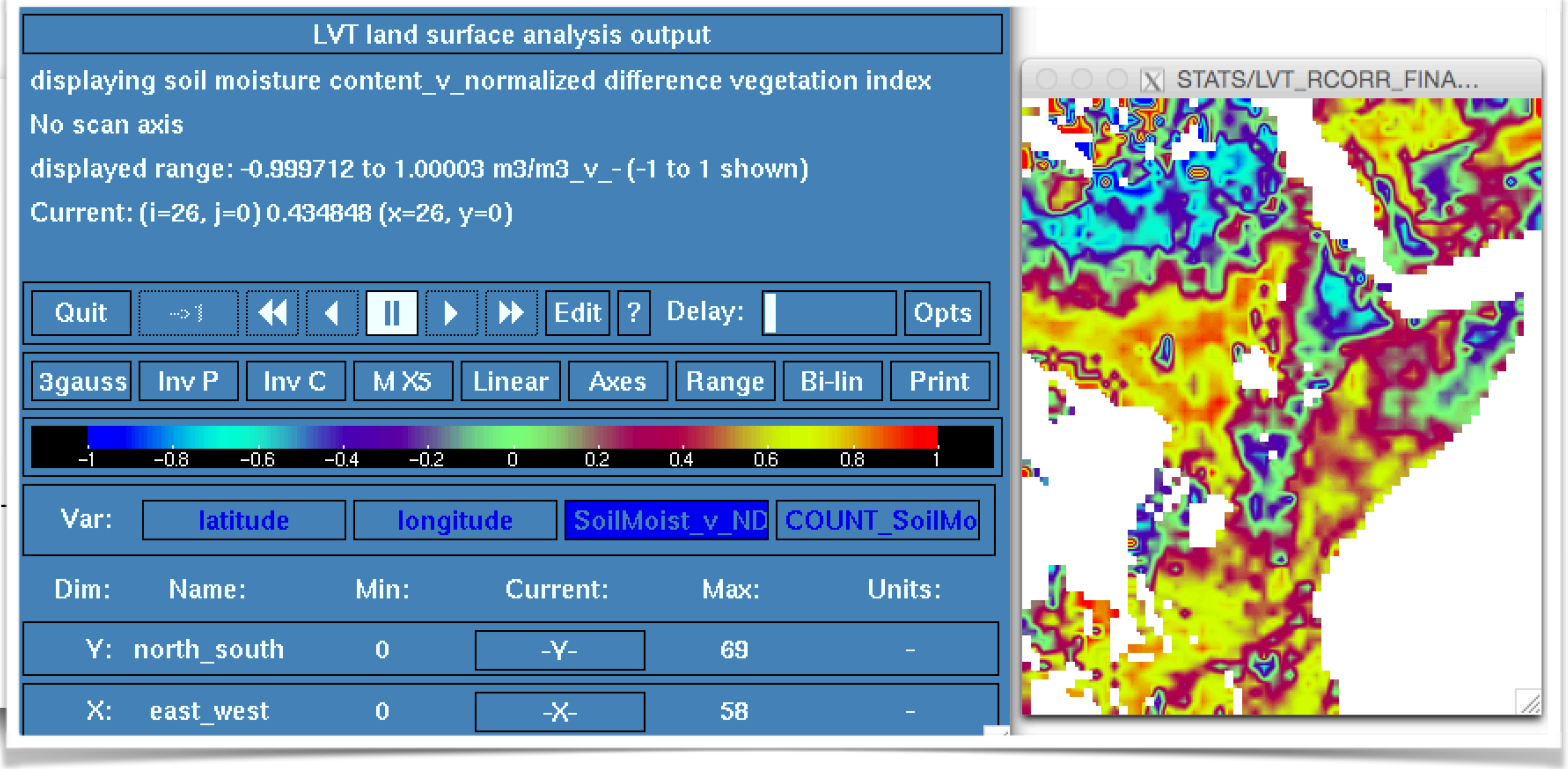


Use a negative temporal lag

```
Temporal lag in metric computations:           "-1mo"
```

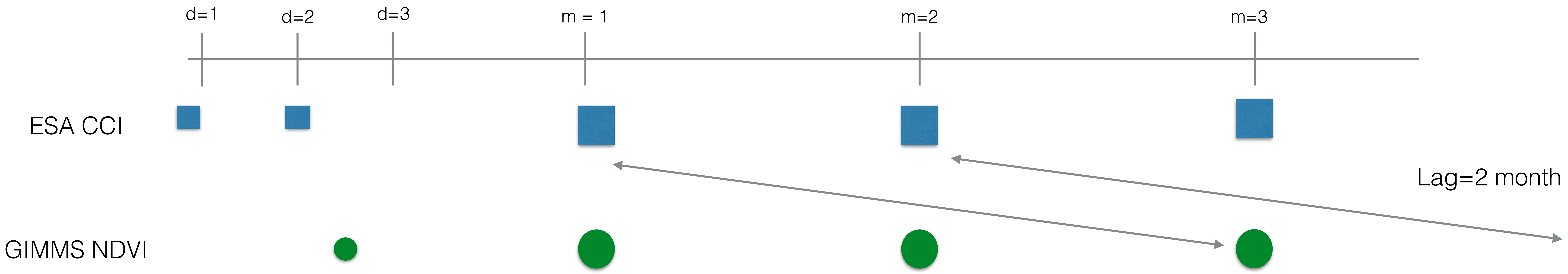



```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % tail lvtlog.0000
 [INFO] Reading ..
 ../DATA_ESACCI/2006/ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED-20061231000000-fv02.0
 .nc
 [INFO] Reading GIMMS NDVI file ../DATA_GIMMSNDVI/2000s_new/geo06nov15a.n17-VI3g
 [INFO] Reading GIMMS NDVI file ../DATA_GIMMSNDVI/2000s_new/geo06nov15b.n17-VI3g
 [INFO] Writing restart file
 ./STATS/RST/LVT.200612310000.rst

 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % more STATS/RCORR_SUMMARY_ST
 ---------------------------------------------------------
 VAR: soil_moisture_content_v_normalized_difference_vegetation_index
 ---------------------------------------------------------
     ALL:       0.251E+00 +/-       0.173E-01           2653
     REG1:      0.251E+00 +/-       0.173E-01           2653
```

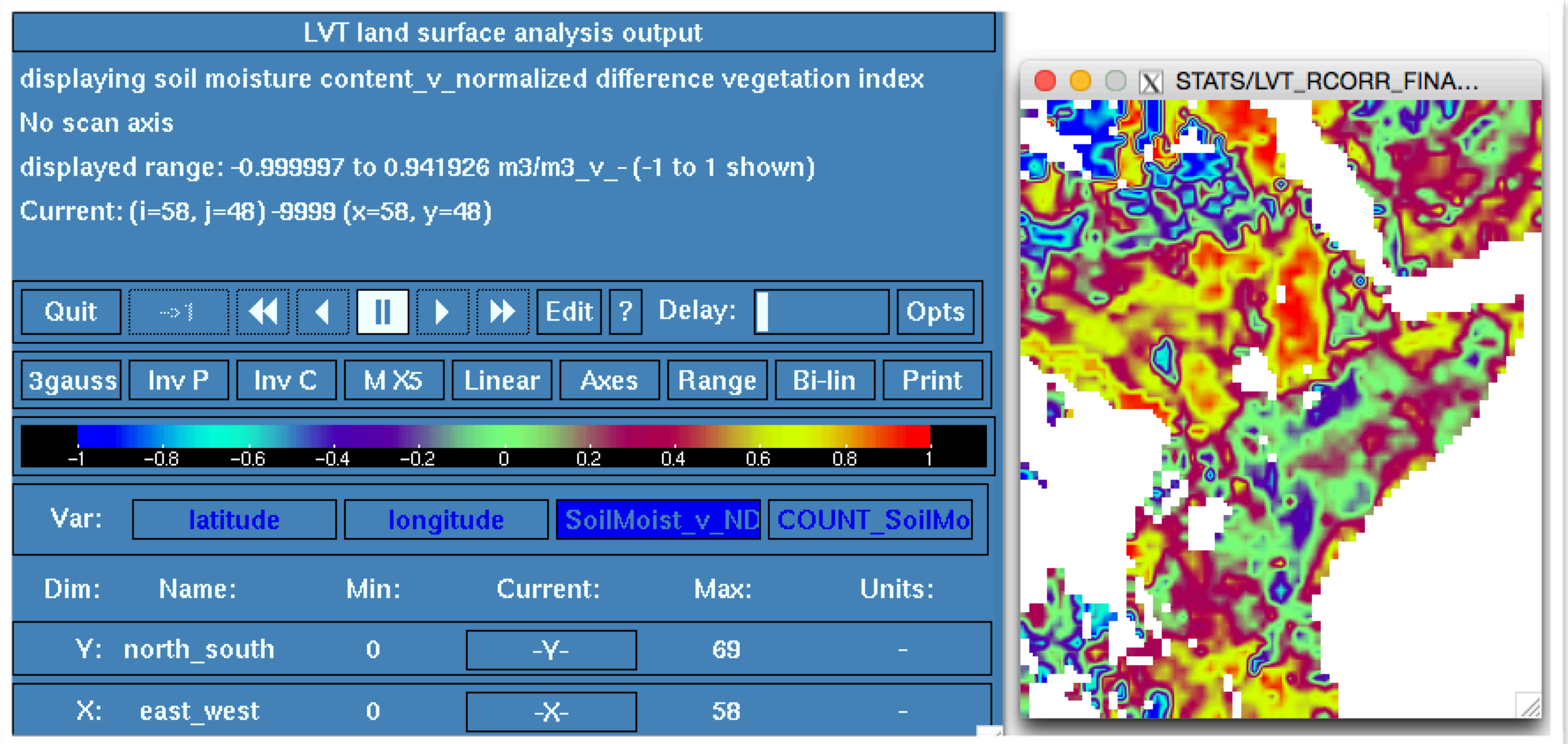


Increase the temporal lag..

```
Temporal lag in metric computations:          "2mo"
```




```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % tail lvtlog.0000
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 12/31/2006 00:00:00
 [INFO] Reading ..
 ../DATA_ESACCI/2006/ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED-20061231000000-fv02.0
 .nc
 [INFO] Writing restart file
 ./STATS/RST/LVT.200612310000.rst

 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/ESACCIvsGIMMSNDVI_ex % more STATS/RCORR_SUMMARY_STATS.dat
 ---------------------------------------------------------
 VAR: soil_moisture_content_v_normalized_difference_vegetation_index
 ---------------------------------------------------------
     ALL:      0.238E+00 +/-       0.162E-01          2659
     REG1:     0.238E+00 +/-       0.162E-01          2659
```

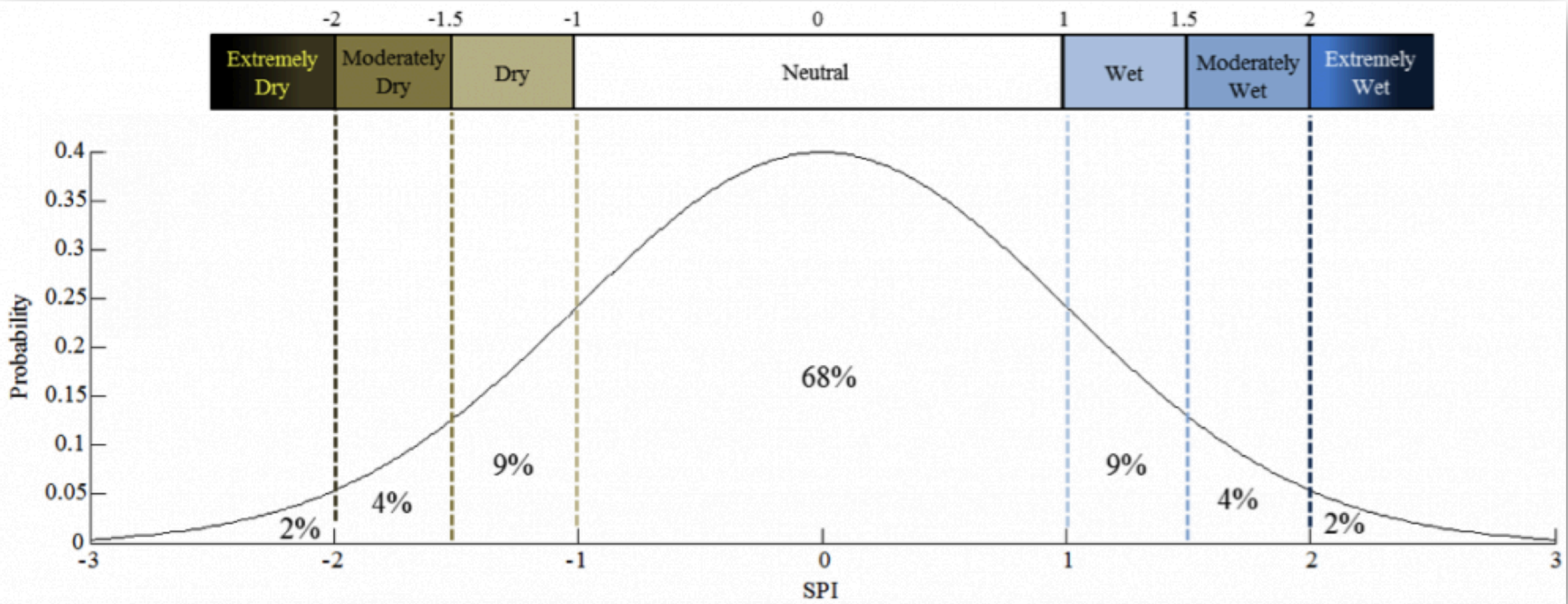


Example 5

Generating drought indicators (SPI)

SPI (standardized precipitation index)

- Widely used as the standard index for quantifying meteorological drought
- Only based on precipitation
- Quantifies observed precipitation as a standardized departure from a selected probability distribution
- Typically precipitation data is fitted a gamma distribution
- Can be interpreted as the number of standard deviations by which the observed anomaly deviates from the long-term mean



SPI labels and their relationship to the normal curve

Category	SPI
Extremely Wet	≥ 2,0
Severely Wet	1,50 to 1,99
Moderately Wet	1 to 1,49
Lightly wet	0 to 0.99
Lightly Drought	0 to -0,99
Moderately Drought	-1 to -1,49
Severely Drought	-1,50 to -1,99
Extremely Drought	≤ -2,0

Example 5 configuration

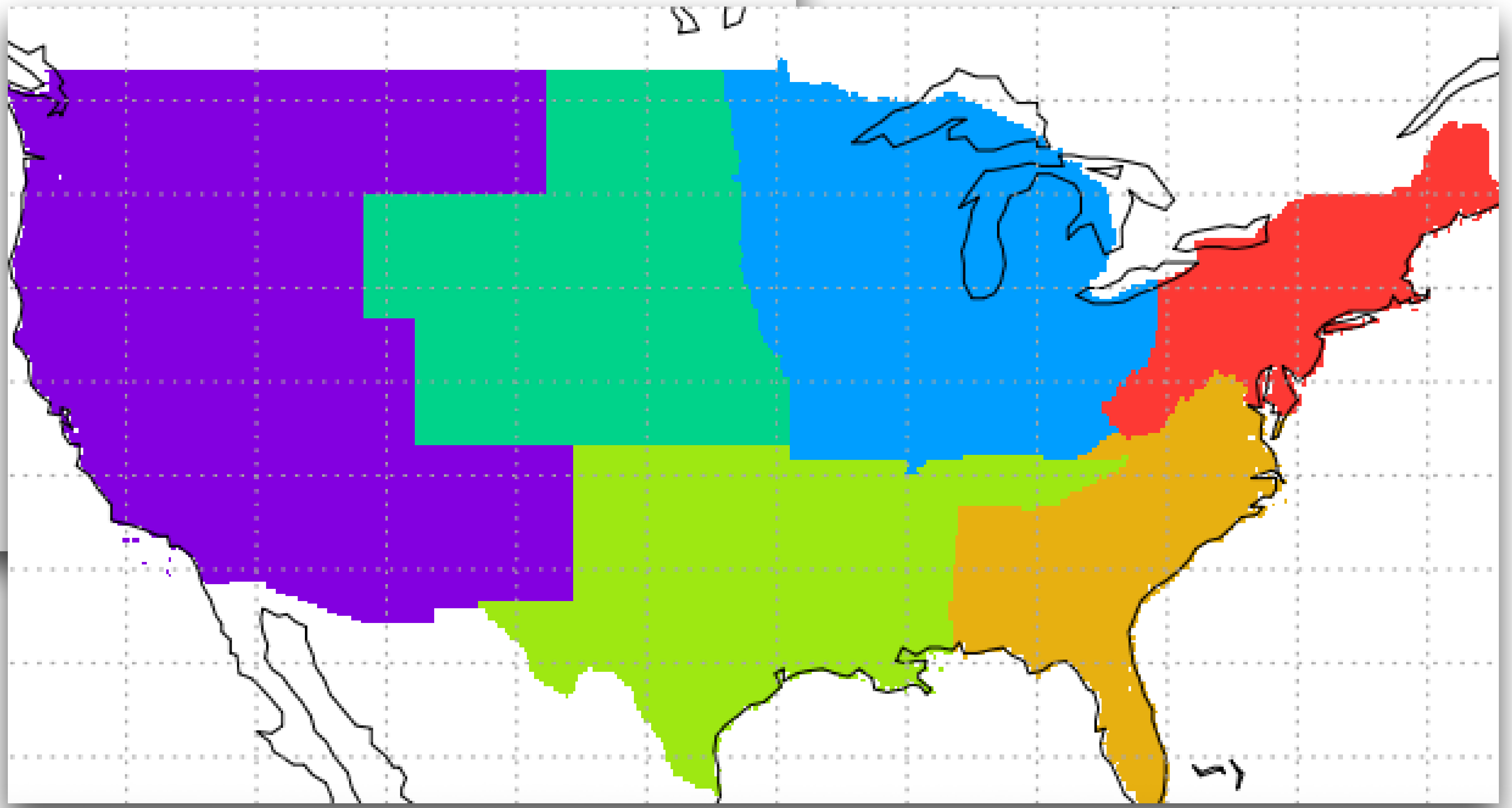
- Computing SPI (and other drought indices) typically requires two passes through the data
  1. First pass to compute the climatology/fit the distribution
  2. Second pass to derive the index relative to the climatology/distribution
- A long archive is desired to ensure enough sampling density in these computations
- Obviously difficult to do all these steps in this tutorial, but LVT also includes the capability to conduct the second step alone (from an already established climatology/distribution), using the restart capabilities.

```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of generating SPI values
# using Rainfall values from a LIS output.
#
# The model output from Noah.3.3 output is produced over CONUS at 0.125 deg
# spatial resolution (at daily intervals).
#
# The simulation is restarted from a previous integration that computed
# the climatology.
#
#------------------------------------------------------------------------------

LVT running mode:          "Data intercomparison"
Map projection of the LVT analysis: "latlon"
LVT output format:         "netcdf"
LVT output methodology:    "2d gridspace"
Analysis data sources:       "LIS output"   "none"

Start mode:                                  restart
LVT restart output interval:                 "1mo"
LVT restart filename:                        LVT.200512310000.rst
Starting year:                               2006
Starting month:                              1
Starting day:                                1
```

The run is being restarted from a previous checkpoint file

Using a categorical map for extracting time series information

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/SPI_ex5 % more TS_LOCATIONS.TXT
#Number of stations
6
#style
5
#names
WEST
1
MIDWEST
1
HIGHPLAINS
1
SOUTH
1
SOUTHEAST
1
NORTHEAST
1
#categorical map
./NLDASmask_USDM.1gd4r
```

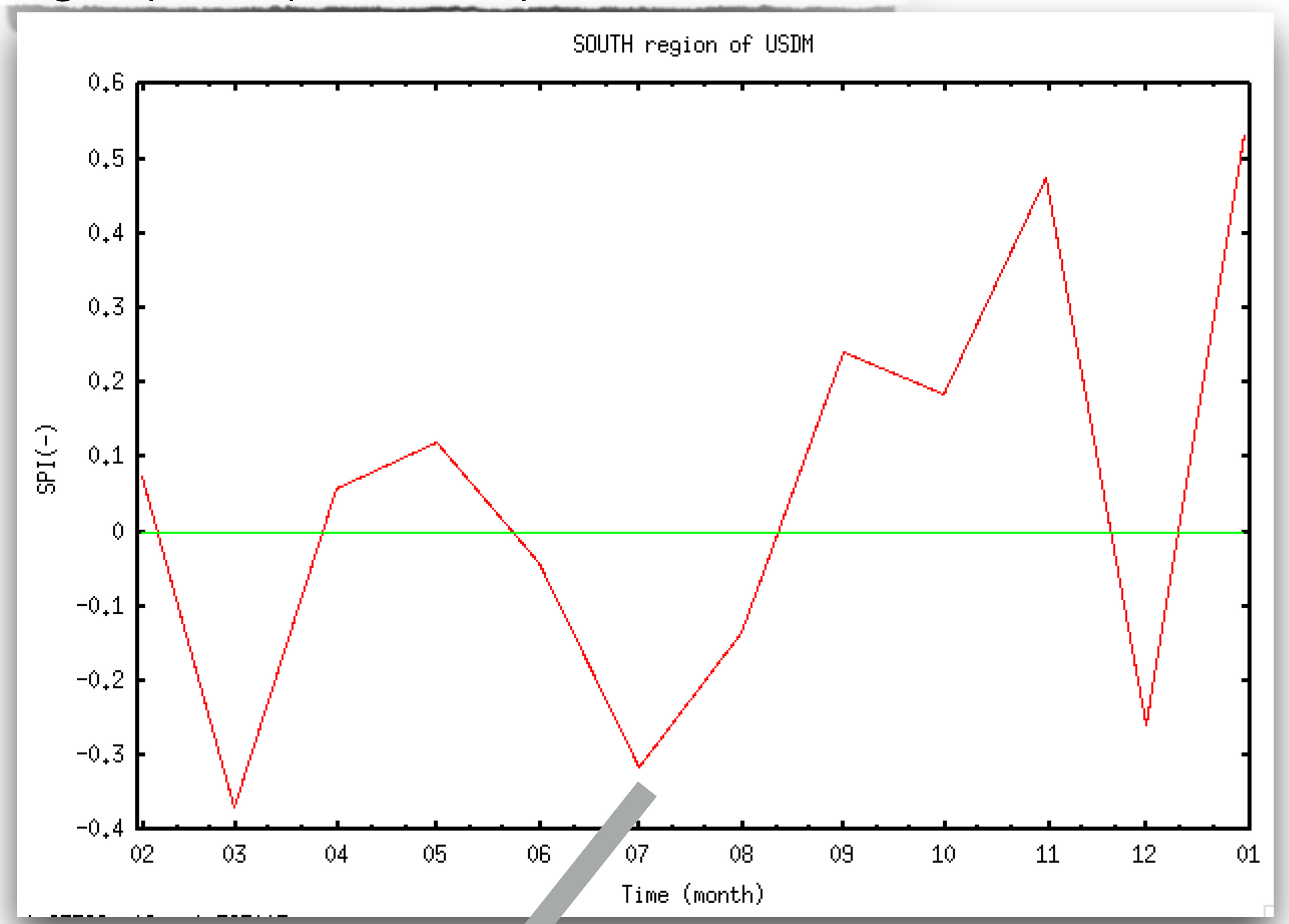


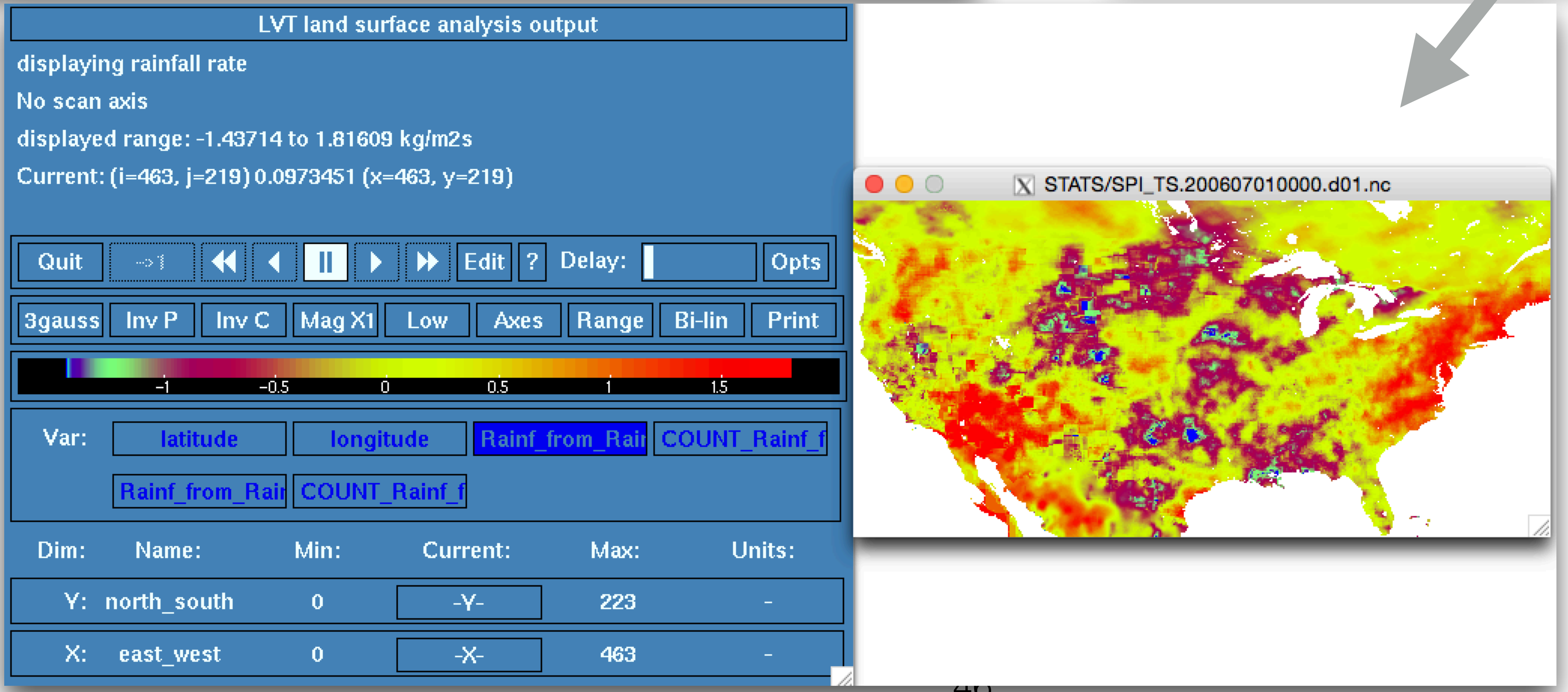
Climatic regions defined by the U.S. Drought Monitor

Examining example 5 output

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/SPI_ex5 % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/SPI_ex5 % tail -20 lvtlog.0000
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612270000.d01.nc
 [INFO] LVT cycle time: 12/28/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612280000.d01.nc
 [INFO] LVT cycle time: 12/29/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612290000.d01.nc
 [INFO] LVT cycle time: 12/30/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612300000.d01.nc
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 12/31/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612310000.d01.nc
 [INFO] Writing restart file
 ./STATS/RST/LVT.200612310000.rst

 [INFO] Finished LVT analysis
 [INFO] ---------------------------------------
```

gnuplot spi_south.plt





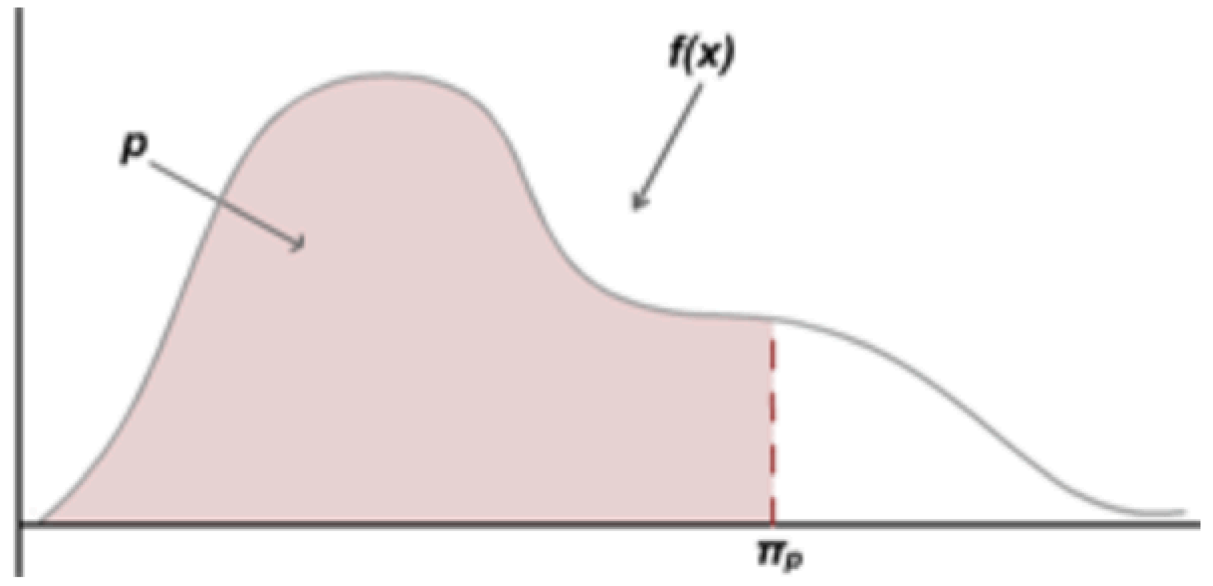
Example 6

Generating drought indicators (Percentiles)

Percentile

- Percentile is a value that indicates the percent of a distribution that is equal to or below it

Definition. If X is a continuous random variable, then the $(\mathbf{100}p)^{th}$ **percentile** is a number π_p such that the area under $f(x)$ and to the left of π_p is p.



That is, p is the integral of $f(x)$ from $-\infty$ to π_p:

$$p = \int_{-\infty}^{\pi_p} f(x)dx = F(\pi_p)$$

Some percentiles are given special names:

- The 25th percentile, $\pi_{0.25}$, is called the **first quartile** (denoted q_1).
- The 50th percentile, $\pi_{0.50}$, is called the **median** (denoted m) or the **second quartile** (denoted q_2).
- The 75th percentile, $\pi_{0.75}$, is called the **third quartile** (denoted q_3).

Source: psu.edu

Percentile metric in LVT

- Percentile metric in LVT can be used with any variable of interest.
- Similar to example 5, computing percentiles requires two passes through the data
  1. First pass to compute the climatology
  2. Second pass to derive the index relative to the climatology/distribution
- LVT employs a modified approach used in the NLDAS drought monitor, where a moving window of 5 days is used to improve the sampling density
- Instead of using a single day across all years, 5 days are used (2 previous days, current day, 2 next days)
  - Jan 3 climatology for example will include Jan 1 - 5 values across all available years
- Not limited to monthly timescales, works with all supported temporal averaging intervals

Example 6 configuration

```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of generating percentile values
# using soil moisture values from a LIS output.
#
# The model output from Noah.3.3 output is produced over CONUS at 0.125 deg
# spatial resolution (at daily intervals).
#
# The simulation is restarted from a previous integration that computed
# the climatology.
#
#------------------------------------------------------------------------------

LVT running mode:          "Data intercomparison"
Map projection of the LVT analysis: "latlon"
LVT output format:         "netcdf"
LVT output methodology:    "2d gridspace"
Analysis data sources:       "LIS output"  "none"

tart mode:                              restart
VT restart output interval:             "1mo"
VT restart filename:                    none
tarting year:                           2006
```

The run is being restarted from a previous checkpoint file, but do not require a restart file for percentile calculations

Instead, the climatology files are expected in the STATS/RST directory (provided with this testcase)

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/Percentile_ex6 % ls STATS/RST/*bin
STATS/RST/001_pctile_climo.bin  STATS/RST/184_pctile_climo.bin
STATS/RST/002_pctile_climo.bin  STATS/RST/185_pctile_climo.bin
STATS/RST/003_pctile_climo.bin  STATS/RST/186_pctile_climo.bin
STATS/RST/004_pctile_climo.bin  STATS/RST/187_pctile_climo.bin
STATS/RST/005_pctile_climo.bin  STATS/RST/188_pctile_climo.bin
STATS/RST/006_pctile_climo.bin  STATS/RST/189_pctile_climo.bin
STATS/RST/007_pctile_climo.bin  STATS/RST/190_pctile_climo.bin
STATS/RST/008_pctile_climo.bin  STATS/RST/191_pctile_climo.bin
```

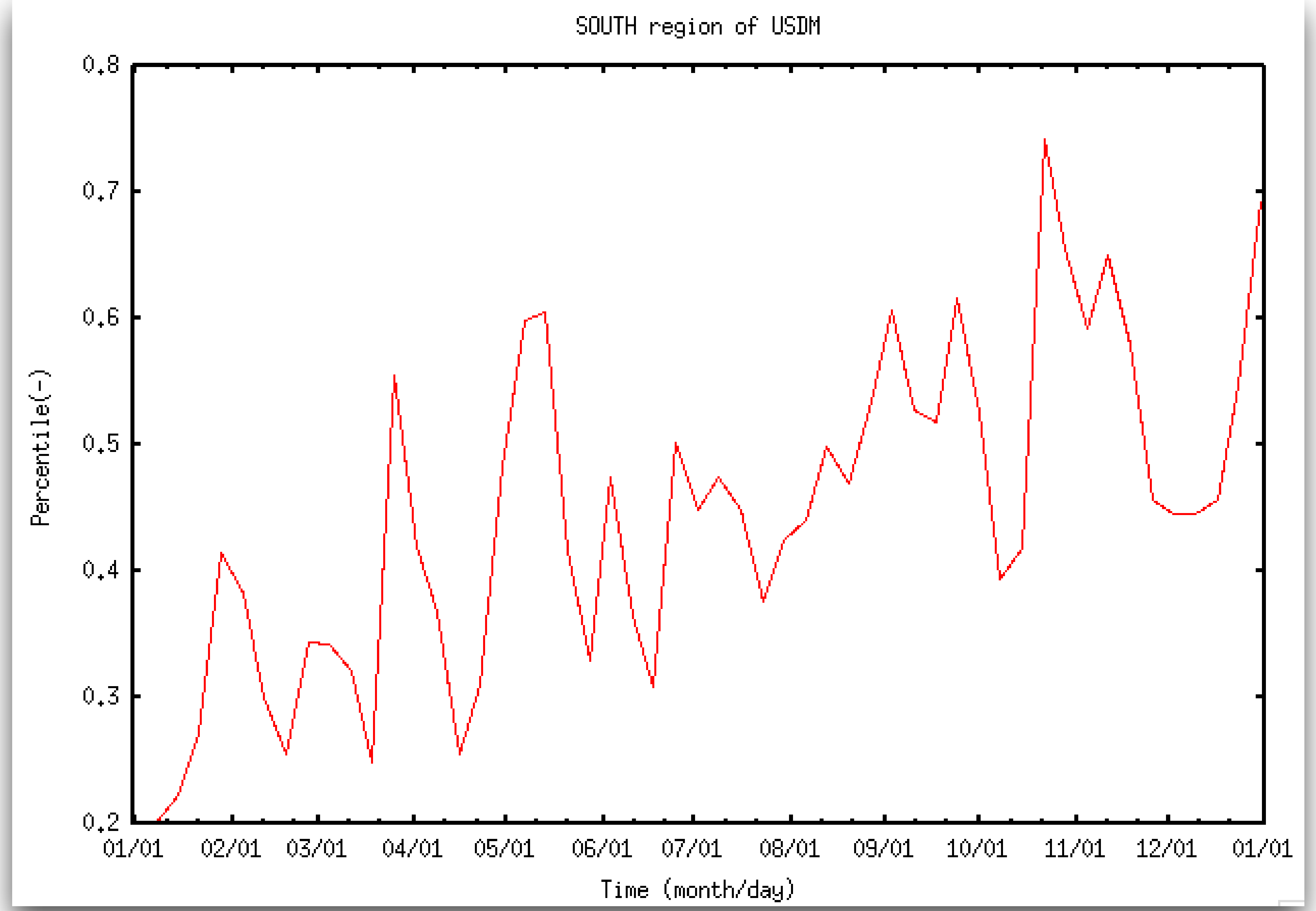
```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/Percentile_ex6 % more METRICS.TBL
#name           total in-time writeTS extractTS threshold SC ADC short_name
Mean:                                  0    0    0    0    -9999.0  0  0  #Mean
Standard deviation:                    0    0    0    0    -9999.0  0  0  #Std
RMSE:                                  0    0    0    0    -9999.0  0  0  #RMSE
Bias:                                  0    0    0    0    -9999.0  0  0  #Bias
ubRMSE:                                0    0    0    0    -9999.0  0  0  #ubRMSE
Mean absolute error:                   0    0    0    0    -9999.0  0  0  #MAE
Anomaly RMSE:                          0    0    0    0    -9999.0  0  0  #ARMSE
Anomaly correlation:                   0    0    0    0    -9999.0  0  0  #ARMSE
Raw correlation:                       0    0    0    0    -9999.0  0  0  #RCORR
Probability of detection (PODy):       0    0    0    0    0.1      0  0  #PODy
Probability of detection (PODn):       0    0    0    0    0.1      0  0  #PODn
False alarm ratio (FAR):               0    0    0    0    0.1      0  0  #FAR
Probability of false detection (POFD): 0    0    0    0    0.1      0  0  #POFD
Critical success index (CSI):          0    0    0    0    0.1      0  0  #CSI
Accuracy measure (ACC):                0    0    0    0    0.1      0  0  #ACC
Frequency bias (FBIAS):                0    0    0    0    0.1      0  0  #FBIAS
Equitable threat score (ETS):          0    0    0    0    0.1      0  0  #ETS
Standard Precipitation Index:          0    0    0    0    -9999.0  0  0  #SPI
Percentile:                            1    1    1    1    -9999.0  0  0  #percentile
```

Percentile calculations are enabled (overall and in time) and the extraction of gridded outputs are also enabled.

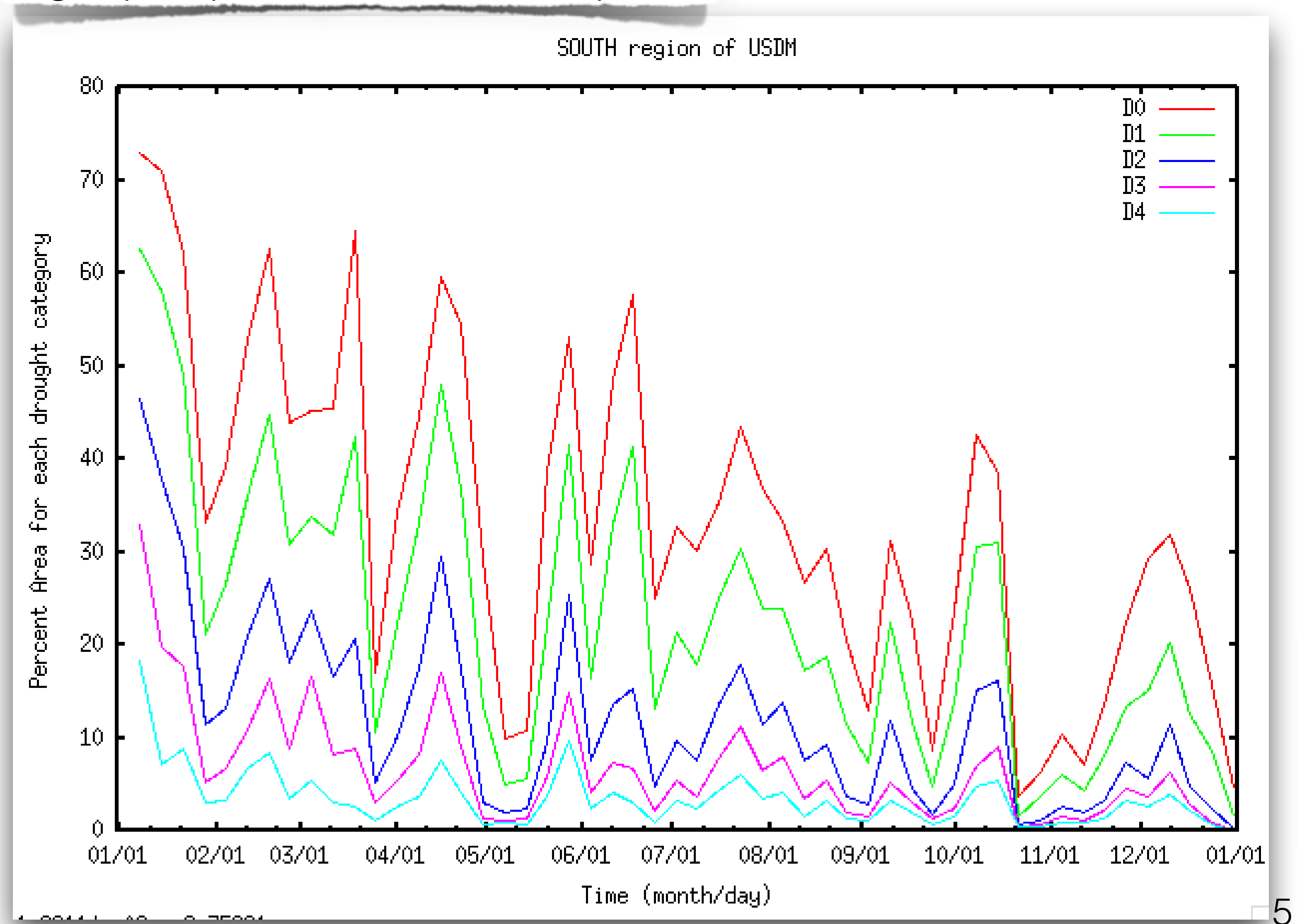
Examining example 6 output

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/Percentile_ex6 % ./LVT lvt.config
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/Percentile_ex6 % tail lvtlog.0000
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612300000.d01.nc
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 12/31/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200612/LIS_HIST_200612310000.d01.nc
 [INFO] Writing restart file
 ./STATS/RST/LVT.200612310000.rst

 [INFO] Finished LVT analysis
 [INFO] ----------------------------------------
```

gnuplot percentile_south.plt



gnuplot percentarea_south.plt



Plots the area under each of the USDM defined drought category (D4 being the most extreme drought)

Example 7

Benchmarking example

- We'll use the in-situ measurements from ARM over SGP to develop a benchmark for latent heat flux (Qle) estimation

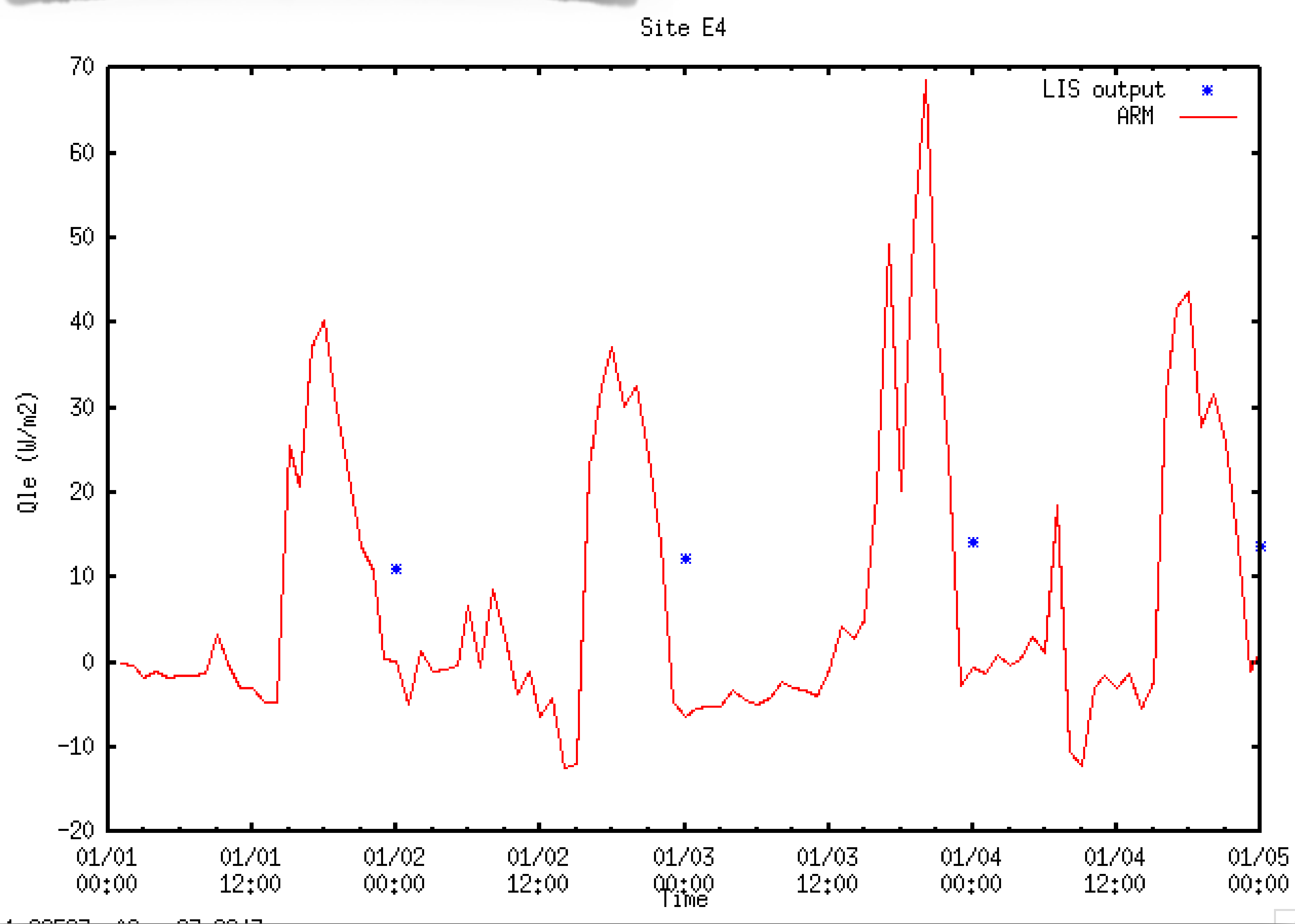
1. Compare the model simulation of Qle vs. ARM data
2. Develop a benchmark using LVT by training ARM Qle measurements to ARM net radiation and air temperature measurements (using out of sample, two-variable regression)
3. Compare the model simulation of Qle vs benchmark

Example 7 : Step 1 (compare model simulation to ARM measurements)

lvt.config_model_arm

```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from a
# LIS output (from Noah.3.3 LSM) against the in-situ ARM-CART
# measurements
#
# The LVT analysis is conducted over a CONUS domain at 0.5 deg
# spatial resolution.
#
# The following variables are compared: Qle
#
# The following metrics are used: Mean
#
#------------------------------------------------------------------------------

LVT running mode:                     "Data intercomparison"
Map projection of the LVT analysis:  "latlon"
Analysis data class:                 "LSM"
LVT output format:                   "netcdf"
LVT output methodology:              "2d gridspace"
Analysis data sources:              "LIS output" "ARM"
```

gnuplot model_arm.plt




```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % ./LVT lvt.config_model_arm
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % tail lvtlog.0000
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE19.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE20.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE22.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE27.b1.20060105.000000.cdf
 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
```

Example 7 : Step 2 configuration

```
#------------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of conducting the
# benchmarking training and generating the benchmark.
#
# The LVT analysis is conducted over a CONUS domain at 0.5 deg
# spatial resolution.
#
# The benchmark is conducted by training the Latent heat fluxes
# against Net radiation and Air temperature from ARM measurements
#
#------------------------------------------------------------------------------
```

```
LVT running mode:                    "Benchmarking"
Map projection of the LVT analysis: "latlon"
LVT output format:                  "netcdf"
LVT output methodology:             "2d gridspace"
Analysis data sources:             "ARM" "ARM"
```

Benchmarking requires a new run mode option

```
LVT datastream attributes table::
Rnet     1   W/m2      -    1   1  Qle 1      W/m2      -    1  1
Tair_f  1    K          -    1   1  Tair_f  0   K          -    1   1
::
```

Datastream attributes table is used to define the training configuration.

Two variables from datastream 1 (Rnet and Tair_f) are used as inputs.

One variable from datastream 2 (Qle) is used as outputs

```
ARM observation directory:              ../DATA_ARM_SGP/   ../DATA_ARM_SGP/
ARM site identifier name:                sgp         sgp
ARM station list file: ../DATA_ARM_SGP/sgp_stns.txt ../DATA_ARM_SGP/sgp_stns.txt
ARM use BAEBBR data:                     1 1
ARM use EBBR data:                       1 1
ARM use ECOR flux data:                  1 1
ARM use SWATS data:                      1 1
ARM use SMOS data:                       1 1
```

Both analysis sources are ‘ARM’. They are specified in consecutive columns

```
Training algorithm for benchmarking:         "Linear regression"
Linear regression mode:                      'two-variable'
Linear regression use out of sample method:  1
```

The output is trained to the inputs using a two-variable linear regression model, using out of sample method

Example 7 : Step 2 (conduct training and generate benchmark)

```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % ./LVT lvt.config_benchmark
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % tail lvtlog.0000
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE19.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE20.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE22.b1.20060105.000000.cdf
 [INFO] Reading ebbr file
 ../DATA_ARM_SGP//2006/sgp30ebbrE27.b1.20060105.000000.cdf
 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % ls STATS.benchmark/
TRAINING/
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % ls STATS.benchmark/TRAINING/
LVT_HIST_OUT_20060101010000.nc  LVT_HIST_OUT_20060102090000.nc  LVT_HIST_OUT_20060103170000.nc
LVT_HIST_OUT_20060101020000.nc  LVT_HIST_OUT_20060102100000.nc  LVT_HIST_OUT_20060103180000.nc
LVT_HIST_OUT_20060101030000.nc  LVT_HIST_OUT_20060102110000.nc  LVT_HIST_OUT_20060103190000.nc
LVT_HIST_OUT_20060101040000.nc  LVT_HIST_OUT_20060102120000.nc  LVT_HIST_OUT_20060103200000.nc
LVT_HIST_OUT_20060101050000.nc  LVT_HIST_OUT_20060102130000.nc  LVT_HIST_OUT_20060103210000.nc
```

```
netcdf LVT_HIST_OUT_20060101020000 {
dimensions:
        east_west = 115 ;
        north_south = 51 ;
        time = 1 ;
variables:
        float latitude(north_south, east_west) ;
                latitude:units = "degree_north" ;
                latitude:standard_name = "latitude" ;
                latitude:long_name = "latitude" ;
                latitude:scale_factor = 1.f ;
                latitude:add_offset = 0.f ;
                latitude:missing_value = -9999.f ;
                latitude:_FillValue = -9999.f ;
        float longitude(north_south, east_west) ;
                longitude:units = "degree_east" ;
                longitude:standard_name = "longitude" ;
                longitude:long_name = "longitude" ;
                longitude:scale_factor = 1.f ;
                longitude:add_offset = 0.f ;
                longitude:missing_value = -9999.f ;
                longitude:_FillValue = -9999.f ;
        float time(time) ;
                time:units = "minutes since 2006-01-01 02:00:00" ;
                time:long_name = "time" ;
                time:time_increment = "1800" ;
                time:begin_date = "20060101" ;
                time:begin_time = "020000" ;
        float Qle(north_south, east_west) ;
                Qle:units = "W/m2" ;
                Qle:standard_name = "surface_upward_latent_heat_flux" ;
                Qle:long_name = "latent heat flux" ;
                Qle:scale_factor = 1.f ;
                Qle:add_offset = 0.f ;
                Qle:missing_value = -9999.f ;
                Qle:_FillValue = -9999.f ;
```

Step 2 produces output files that includes the outputs of Qle generated through the trained model

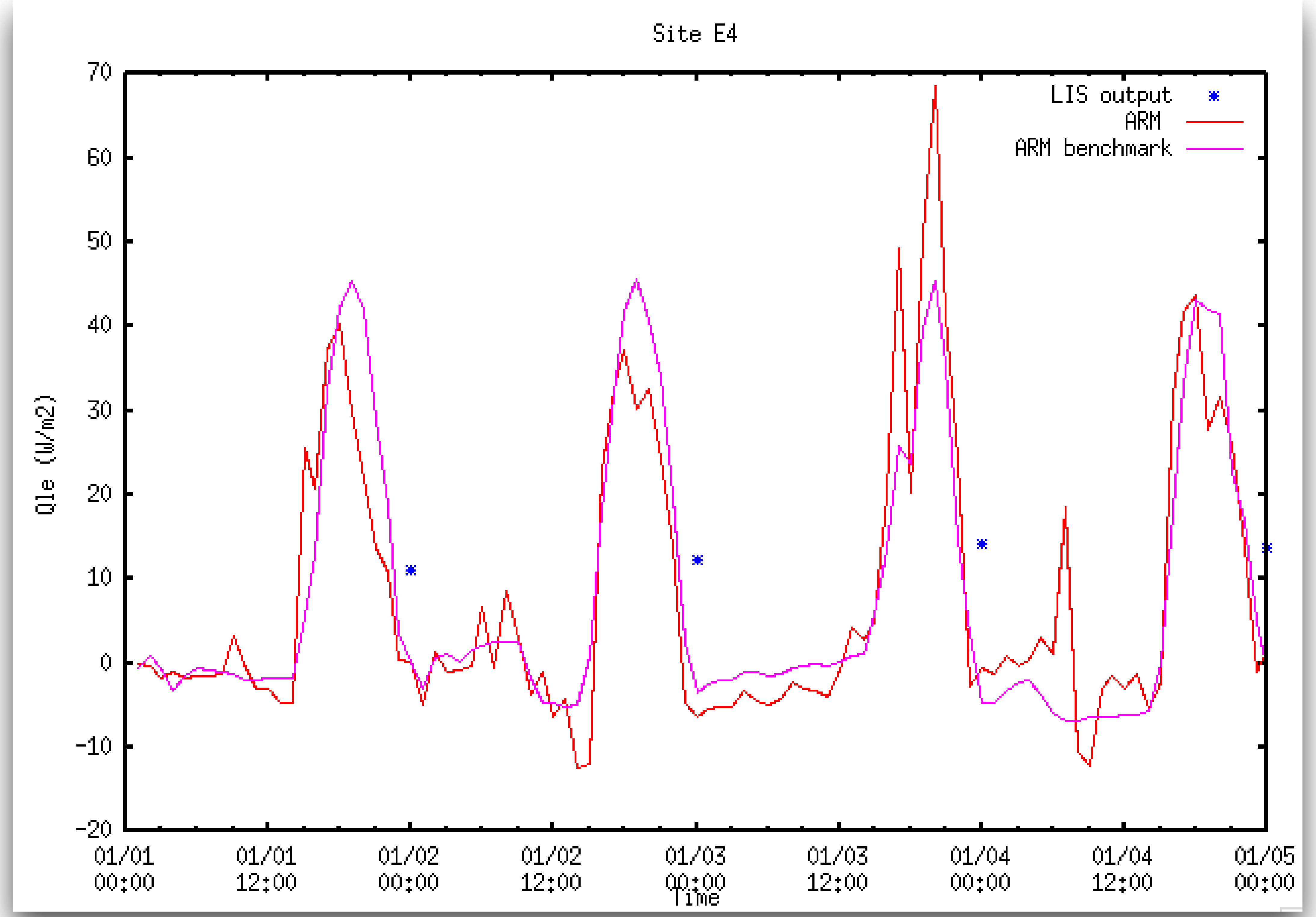
Example 7 : Step 3 (compare model simulation to benchmark)

lvt.config_model_benchmark

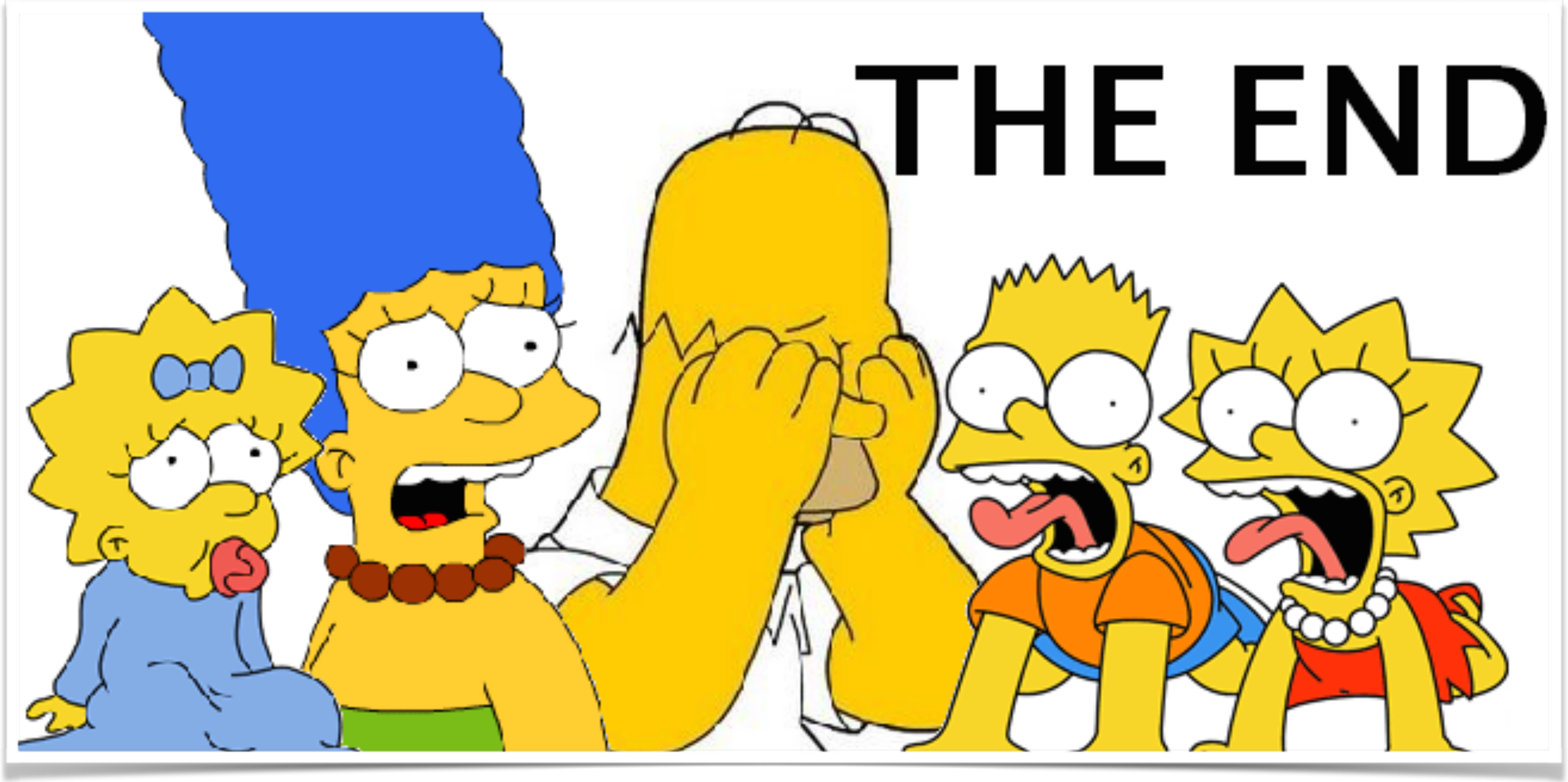
```
#-------------------------------------------------------------------------
# README
#
# This LVT configuration shows an example of comparing variables from a
# LIS output (from Noah.3.3 LSM) against the benchmark output from LVT
#
# The LVT analysis is conducted over a CONUS domain at 0.5 deg
# spatial resolution.
#
# The following variables are compared: Qle
#
# The following metrics are used: Mean
#
#-------------------------------------------------------------------------

LVT running mode:                    "Data intercomparison"
Map projection of the LVT analysis: "latlon"
Analysis data class:                "LSM"
LVT output format:                  "netcdf"
LVT output methodology:             "2d gridspace"
Analysis data sources:             "LIS output" "LVT benchmark"


LVT benchmark output directory:          'STATS.benchmark'
LVT benchmark variable:                   Qle
```

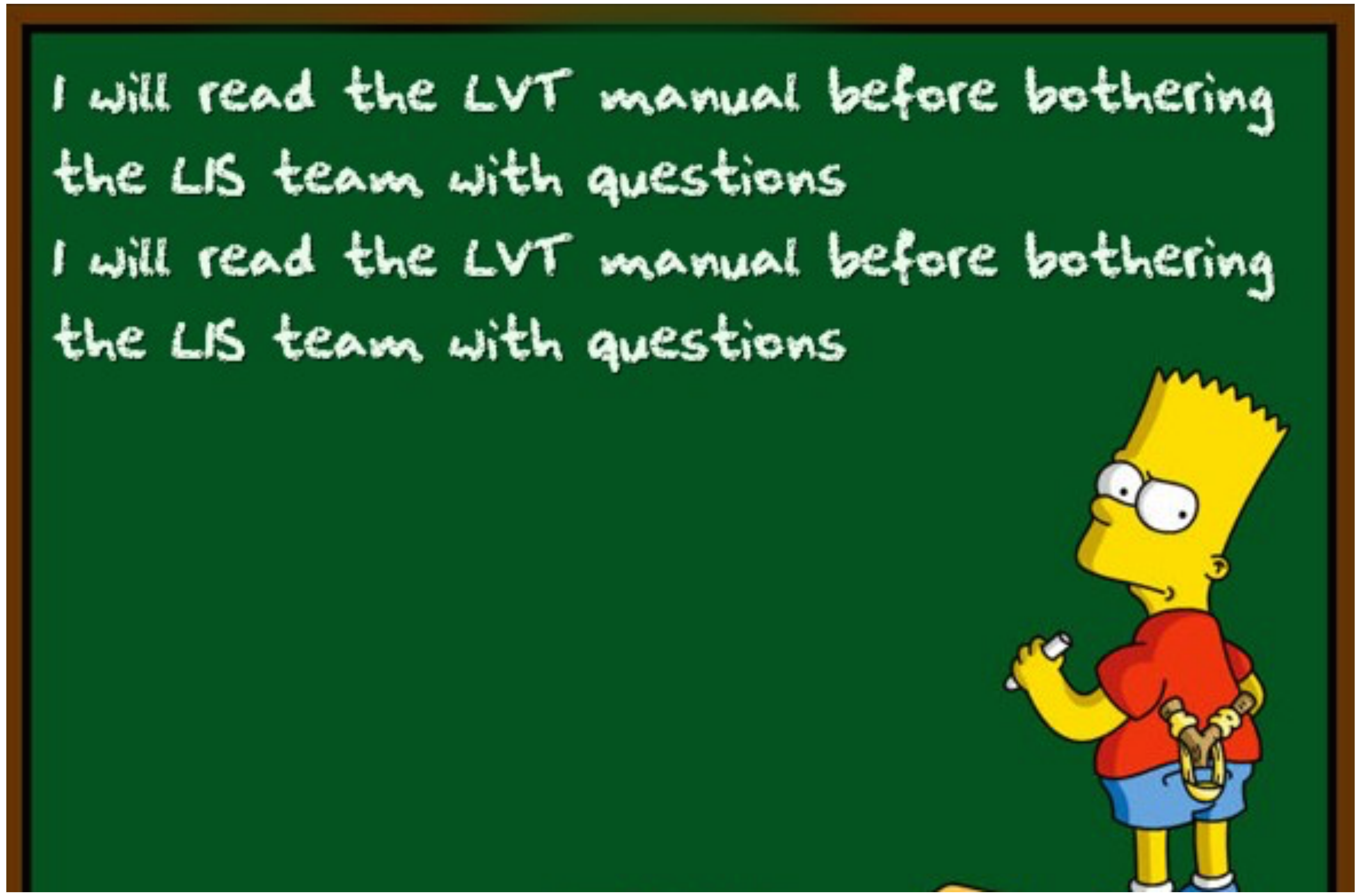
```
/discover/nobackup/projects/lis/Projects/LVT/Tutorial/BenchMark_ARM_ex7 % tail lvtlog.0000
 [WARN] Warning: LVT file
 STATS.benchmark/TRAINING/LVT_HIST_OUT_20060104233000.nc does not exist
 [INFO] LVT cycle completed
 [INFO] LVT cycle time: 01/05/2006 00:00:00
 [INFO] Reading LIS output
 ../DATA_Noah33_CONUS/OUTPUT/SURFACEMODEL/200601/LIS_HIST_200601050000.d01.nc
 [INFO] reading LVT output
 STATS.benchmark/TRAINING/LVT_HIST_OUT_20060105000000.nc
 [INFO] Finished LVT analysis
 [INFO] --------------------------------------
```

gnuplot model_benchmark_arm.plt



Probably not the cleanest/fairest example. But the regression model does a reasonable job of capturing the Qle estimates





Questions/Comments/Feedback?